SEARCH REQUEST FORM

Scientific and Technical Information Center

Access DB# 86576

Requester's Full Name: _Celine Qian_   Examiner # : _75710_ Date: _2/12/03_
Art Unit: _1636_   Phone Number 30_6-0283_   Serial Number: _09/755962_
Mail Box and Bldg/Room Location: _____   Results Format Preferred (circle): PAPER DISK E-MAIL

If mor than one search is submitted, please prioritize searches in order of need.
****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _Expression of Foreign genes from plant virus vectors_
Inventors (please provide full names): _Simon Santa-Cruz et al._

Earliest Priority Filing Date: _1/7/2001_

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search SEQ ID NO: 1, and also do a nucleotide search.

oligo

NA
1-188

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

BEST AVAILABLE COPY

****

STAFF USE ONLY

Type of Search	Vendors and cost where applicable
NA Sequence (#) _1-1_	STN _Review to AA_
AA Sequence (#) ____	Dialog ____
Structure (#) ____	Questel/Orbit ____
Bibliographic ____	Dr.Link ____
Litigation ____	Lexis/Nexis ____
Fulltext ____	Sequence Systems _$4_
Patent Family ____	WWW/Internet ____
Other ____	Other (specify) ____

Searcher: _____
Searcher Phone #: _____
Searcher Location: _____
Date Searcher Picked Up: _2/13/03_
Date Completed: _2/19/03_
Searcher Prep & Review Time: _____
Clerical Prep Time: _____
Online Time: _____

PTO-1590 (8-01)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: February 19, 2003, 06:08:11 ; Search time 2601 seconds
(without alignments)
2103.548 Million cell updates/sec

Title: US-09-758-962-1
Perfect score: 188
Sequence: 1 gtcgacggtatcgataagct.........gaagaaggagctcaccatg 188

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 1455140287 residues
Word size :   0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*

```
1:  gb_ba:*
2:  gb_btg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_v1:*
16: gb_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_v1:*
31: em_htg_hum:*
32: em_htg_inv:*
33: em_htg_other:*
34: em_htg_mus:*
35: em_htg_pln:*
36: em_htg_rod:*
37: em_htg_man:*
38: em_htg_vrt:*
39: em_sy:*
40: em_htgo_hum:*
41: em_htgo_mus:*
    em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	149	79.3	1537	14	S68164	S68164 transport p
2	149	79.3	6312	14	TMVRPTPCP	Z29370 Tobacco mos
3	42	22.3	6311	14	BRU03387	U03387 Turnip vein
4	38	20.2	1568	14	MDVPRA	L41561 Galiid herp
5	35	18.6	2126	6	AR067739	AR067739 Sequence
6	35	18.6	2126	6	AR096004	AR096004 Sequence
7	35	18.6	2126	6	AR116595	AR116595 Sequence
8	35	18.6	2126	6	AR121154	AR121154 Sequence
9	35	18.6	7560	6	AR153837	AR153837 Sequence
10	34	18.6	110	10	AB027145	AB027145 Rattus no
11	34	18.1	338	11	CSPAG5R	Z47159 Calothrix D
12	34	18.1	369	11	G16087	G16087 chicken STS
13	34	18.1	373	11	G01691	G01691 chicken STS
14	34	18.1	397	1	CSPAG13R	Z47173 Calothrix D
15	34	18.1	487	6	HRSTCRAG	L27850 Equus cabal
16	34	18.1	741	6	AX156405	AX156405 Sequence
17	34	18.1	759	6	E14482	E14482 cDNA encodi
18	34	18.1	924	8	A59874	A59874 Sequence 3
19	34	18.1	924	6	BNPBP3	X90729 B.napus mRN
20	34	18.1	1043	8	CAR011383	AJ011383 Cicer ari
21	34	18.1	1097	8	AB049721	AB049721 Pisum sat
22	34	18.1	1181	10	E14788	E14788 cDNA encodi
23	34	18.1	1221	10	MMIGEBP1	X55499 M.musculus
24	34	18.1	1248	6	AX304651	AX304651 Sequence
25	34	18.1	1318	6	E37094	E37094 Sequence
26	34	18.1	1438	9	AB067774	AB067774 Homo sapi
27	34	18.1	1464	6	AX286689	AX286689 Sequence
28	34	18.1	1638	6	AR107064	AR107064 Sequence
29	34	18.1	1638	6	AR136732	AR136732 Sequence
30	34	18.1	1638	6	AR169244	AR169244 Sequence
31	34	18.1	1680	6	XLU09632	U09632 Xenopus lae
32	34	18.1	1845	8	AB085846	AB085846 Chlorella
33	34	18.1	1889	6	AX049343	AX049343 Sequence
34	34	18.1	1889	6	A94995	A94995 Sequence 1
35	34	18.1	1958	6	AX022639	AX022639 Sequence
36	34	18.1	1958	6	E38366	E38366 Binding par
37	34	18.1	1958	8	E39010	E39010 Nucleic aci
38	34	18.1	1990	8	CCA401353	AJ401353 Cuscuta c
39	34	18.1	1998	8	CPL401355	AJ401355 Cuscuta P
40	34	18.1	2142	8	SOJ002265	AJ002265 Spinacia
41	34	18.1	2171	4	AB051103	AB051103 Felis cat
42	34	18.1	2185	6	AR092424	AR092424 Sequence
43	34	18.1	2199	3	LME292039	AJ292039 Leishmani
44	34	18.1	2200	6	AR050076	AR050076 Crucifer
45	34	18.1	2378	14	AB003936	AB003936 Crucifer

ALIGNMENTS

```
RESULT 1
S68164
LOCUS       S68164               1537 bp    DNA     linear   VRL 23-SEP-1994
DEFINITION  transport protein, capsid protein [Tobamovirus, host: cruciferous
            plants, Genomic, 1537 nt].
ACCESSION   S68164
VERSION     S68164.1  GI:544600
KEYWORDS    .
SOURCE      Tobamovirus host: cruciferous plants.
  ORGANISM  Tobamovirus
            Viruses; ssRNA positive-strand viruses, no DNA stage.
REFERENCE   1  (bases 1 to 1537)
  AUTHORS   Dorokhov,Iu.L., Ivanov,P.A., Novikov,V.K., Yefimov,V.A. and
            Atabekov,I.G.
  TITLE     Tobamovirus of cruciferous plants: nucleotide sequence of genes of
```

```
RESULT 2
TMVRPTPCP
LOCUS       TMVRPTPCP          6312 bp    RNA    linear   VRL 16-SEP-1997
DEFINITION  Tobacco mosaic virus (Crucifer) genomic RNA for RNA-dependent RNA
            polymerase; 122k protein; transport protein and coat protein.
ACCESSION   Z29370
VERSION     Z29370.1  GI:488713
KEYWORDS    122K protein; coat protein; RNA helicase; RNA-dependent RNA
            polymerase; transport protein.
SOURCE      Tobacco mosaic virus.
  ORGANISM  Tobacco mosaic virus.
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1 (bases 1 to 6311)
```

```
Qy     27  CGAATTCGTCGATTCGGTTCGCAGCACATTTAAAGCCGGTTGACAACTTTAAAAGAAGAAAA  86
           |||||||||||||||||||| |||||| |||||||||||||||||||||||||||||||
Db    681  CGAATTCGTCGATTCGGTTGCAGCATTTAAAGCCGGTTGACAACTTTAAAAGAAGAAAA  740

Qy     87  GAAGGTTGAAGAAAAAGGGTGTAGTAAGTAAGTATAGAGACGAGGAAGTACGCCGG  146
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741  GAAGGTTGAAGAAAAAGGGTGTAGTAAGTAAGTATAGAGACGAGGAAGTACGCCGG  800

Qy    147  TCCTGATTCGTTTAATTTGAAAGAAGAAA  175
           |||||||||||||||||||||||||||||
Db    801  TCCTGATTCGTTTAATTTGAAAGAAGAAA  829
```

```
Query Match   79.3%;   Score 149;   DB 14;   Length 1537;
Best Local Similarity   100.0%;   Pred. No. 1.6e-74;
Matches 149;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

ORIGIN
BASE COUNT      423 a    274 c    406 g    433 t    1 others
```

```
FEATURES        Location/Qualifiers
     source  1..1537
                     /organism="Tobamovirus"
                     /db_xref="taxon:12234"
     gene    103..906
                     /gene="transport protein"
     CDS     103..906
                     /product="transport protein"
                     /protein_id="AAB29319.1"
                     /db_xref="GI:544601"
                     /translation="MSIVSYEPKVSDFLNLSKKEEILPKALTLRLKTVSISTKDIISVK
                     ESETLCDLIINVPLDKYRYVGILGAVFTGEWLVPDFVKGGVTISVIDKRLANSKEC
                     VIGTYRAAAKSRKFQFKLVPNYFVSTVDARKPWQVHVRIQDLKIEAGWQPLALEVVS
                     VAMVTNNVVMKGLREKVVAINDPDVEGFEGVVDEFVDSVAAFKAVDNFRRKKVEEK
                     GVSKYKYRPEKYAGPDSFNLKEENVLQHYKPESVPFRSGVGRAHSDA"
     gene    830..1219
                     /gene="capsid protein"
     CDS     830..1219
                     /gene="capsid protein"
                     /note="This sequence comes from Fig. 2"
                     /product="capsid protein"
                     /codon_start=1
                     /protein_id="AAB29320.1"
                     /db_xref="GI:544602"
                     /translation="MSYNITNPNQYQYFAAVWAEPIPMLNQCISALSQSYQTQAARDT
                     VRQFSNLLSAVVAPSQRPETGSRVYVNSAVIKPLYEALMKSPDTRNRIIETEEESR
                     PSASEVLTRHNVLMRPPLEVKPSCC"
```

```
REMARK      GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 142918] from the original journal article.
            This sequence comes from Fig. 2.
FEATURES        Location/Qualifiers
     source  1..1537
                     /organism="Tobamovirus"
                     /db_xref="taxon:12234"
     gene    103..906
                     /gene="transport protein"
     CDS     103..906
                     /product="transport protein"
                     /note="This sequence comes from Fig. 2; Protein sequence
                     is in conflict with the conceptual translation;
                     mismatch(95[R->G])"
                     /codon_start=1
```

```
AUTHORS     Dorokhov,Y.L., Ivanov,P.A., Novikov,V.K., Yefimov,V.A. and
            Atabekov,I.G.
TITLE       Tobamovirus of cruciferous plants: nucleotide sequence of genes of
            the transport protein, capsid protein, and 3'-terminal untranslated
            region
JOURNAL     Dokl. Akad. Nauk. 332 (4), 518-522 (1993)
MEDLINE     94083923
PUBMED      8260923
JOURNAL     the transport protein, capsid protein, and 3'-terminal untranslated
            region
```

```
AUTHORS     Dorokhov,Y.L., Ivanov,P.A., Novikov,V.K., Yefimov,V.A. and
            Atabekov,I.G.
TITLE       Tobamovirus of cruciferous plants: nucleotide sequence of genes of
            the transport protein, capsid protein, and 3'-terminal untranslated
            region
JOURNAL     Dokl. Akad. Nauk SSSR 332, 518-522 (1993)
REFERENCE   2 (bases 1 to 6312)
AUTHORS     Dorokhov,Y.
            Direct Submission
JOURNAL     Submitted (13-JAN-1994) Dorokhov Y., A.N.Belozersky Institute,
            Moscow State University, Moscow, Russia
REMARK      revised by author 15-MAR-94
REFERENCE   3 (bases 1 to 6312)
AUTHORS     Morozov,S.Yu., Efimov,V.A., Novikov,V.K., Agranovsky,A.A.,
            Morozov,Yu.L., Ivanov,P.A., Casper,R. and Atabekov,J.G.
TITLE       Complete nucleotide sequence and genome organization of a
            tobamovirus infecting cruciferous plants
JOURNAL     FEBS Lett. 350 (1), 5-8 (1994)
MEDLINE     94341372
PUBMED      7545946
REFERENCE   4 (bases 1 to 6312)
AUTHORS     Dorokhov,Y.L., Ivanov,P.A., Novikov,V.K., Efimov,V.A.,
            Dorokhov,Y.L., Efimov,V.A., Casper,R. and Atabekov,J.G.
TITLE       Tobamovirus of the Cruciferae family; nucleotide sequence of the
            obamovirus of the Cruciferae family; nucleotide sequence of the
            5'-untranslated region and nonstructural protein genes controlling
            replication viral genome
JOURNAL     Dokl. Akad. Nauk SSSR 335, 792-798 (1994)
REFERENCE   5 (bases 1 to 6312)
AUTHORS     Belenovich,E., Generozov,E., Novikov,V. and Zavriev,S.
TITLE       Properties and the genome structure of the K2 strain of tobacco
            mosaic virus
JOURNAL     Mol. Biol. (Mosk.)
FEATURES        Location/Qualifiers
     source  1..6312
                     /organism="Tobacco mosaic virus"
                     /db_xref="taxon:12242"
     misc_feature   1..6312
                     /note="genomic RNA of the tobacco mosaic virus infecting
                     cruciferous plants"
                     /protein_id="CAA82559.1"
                     /db_xref="GI:619908"
     CDS     69..4874
                     /function="protein containing RNA-dependant RNA
                     polymerase"
                     /note="ORF2; readthrough protein"
                     /codon_start=1
                     /transl_except=(pos:3390..3392,aa:OTHER)
                     /product="178K"
                     /protein_id="CAA82558.1"
                     /db_xref="SPTREMBL:Q88604"
                     /translation="MAQFQQTIDMQTLQQAAGPNSLVNDLASRRVVDNAVEELNARSR
                     RPKVHFSKAVSTEQTLIATNAYPEFEISFTHTQSAVHSLAGGFRSLELEYLMMQVPEG
                     SLTYDIGGNESAHLFKGRDYVHCCMPNLDVRDIARHEGHKEAIHSYVNRLKRQQRPVP
                     EYQRAAFNNYAENPHFVHCDKPFQQCELTTANGTDYAVALLSIYIDIPVEEFGSALLR
                     KNVKTCFAAFHFHENNLLDCDTVTLDEIGATFQRAGNLSFFHNESTLNYTHSFSNI
                     IKYVCKTFFPASQRFVYHKEFLVTRVNTWYCKFTRVDTFTLFRGVIHNNVDCEEFYKA
                     MDDAWHYKKTLAMLNAERTIPKDNAALNFWFPKVRDMVIVPLFDASITTGRMSRREVM
                     VNKDFVYTVLNHIKTYQAKALTYANVLSFVESIRSRVIINGVTARSEWDTDKAILGPL
                     AMTFLIITKLGHVDDEILKRFQKFDRTTNELIWTSLCDALMGVIPSVKETLVRGGFV
                     KVAEQANLEIKIPELCYCTFADRLVLQYKKAEEFQSCDLSKFEEESEKYYNALSELSVLE
                     SAEHKDVLSLQNDAPFCVKNLVEGSVPAYGMCPKGGGFDKDLDVDIADFHLKSVDAVK
                     RGTMMSAVYTGSSIEVRQMKNYIDYLSALSATSVSNICKVLRDLGDPOSEQEKSGVWD
                     VRGRWLLRPNAKSHAWGVAEDANHKLVIVLLNWDDGKPVCDETWRVAVSSDSLIYS
                     DMGKLKTLTCSPNGEPPENAKVILVDGVPGCGKTKEIIEKVNFSEDLLVPGKEAS
                     KMIIRRANHAGVIRADKDNVSTVDSFLMHPSRVRFLDGELMHTGCVNFLLLS
                     QCDVAVYYGDTQQIFPICRVANFPYPAHFAKLVADEKEVRRVTLRCPADYTYFLNKKY
                     DGAVMCTSAVERSRKAEVVRGKGALNPITLPLEGKILTFTQQADKFELLEGKYKDVNTV
                     HEVQGETYERTAIVRLTPLEIISRASPHVLVALTRHTRCKYTVVLDPMVNISE
                     MEKLSNFLLDMYRVEAGIQXOLQIDAVFKGTNLFVQTPKSGDWRDMQFYNDTLLPQNS
                     TILNEVDAVTMNLRDISLNVKDCRIDFSKSVQLPKEQPILKPKIRTAAEMPRTAGLL
                     ENLVAMIKRNMNAPDLTGTIDIEDTASLVVEKFMDSYIDKEFSGGTNEMTMTRESFSRW
                     LSKQESSTVGQLADFNFVDLPAVDEYKHMIKSQPKQLIDLSIQDEYPALQCIVYHSKK
                     INAIFGPMFSELTRMLLERIDSSKFLFYTRKTPAQOIEDFESDLDSTQAMEILEDISK"
```

us-09-758-962-1.rge

Left column:

```
     CDS
                     YDKSQNEFHCAVEYKIWEKLGIDEWLAEVWKQGHRKTTLKDYTAGIKTCLWYQRKSGD
                     VVTTFIGNTIIIAACLSSMIPMDKVIKAAPCGDDSLIYIPKGLDLPIOAGANLMWNFE
                     AKLFRKKYGFGCGRYVIHHDRGAIVYDGRLLISKLGCKHIRDVHLEELRESLCDVA
                     SNLNNCAYFSQLDEAVAEVHKTAVGGSFAFCSIIKYLSDKRLFRDLFYV"
     69. .3392
                     /codon_start=1
                     /note="ORF1"
                     /product="122K"
                     /protein_id="CAA82558.1"
                     /db_xref="GI:488714"
                     /db_xref="SPTREMBL:Q88603"
                     RNA-helicase domains"
                     /function="protein containing methyltransferase &
                     /translation="MAQFQQTIDMQTLQAAAGPNSLVNDLASRRVYDNAVEELNARSR
                     EYQRAAFNNYAENPHFVHCDKPFFQQCELTPANGTDIYAVALLSIYDIPVEEFGSALLR
                     RPKVHFSKAVSTEQTLIATNAYPEFEISFTHTQSAVHSLAGGFRSLELEYLMMQVPFG
                     SLTYDIGGNFSAHLFKGRDYVHCCMPNLDVRDIARHEGHKEAIHSYVNRLKRQQRPVP
                     KNVKTCFAAFHFHENMLLDCDTVTLDEIGAFQKRAGDNLSFFHNESTLNYTHSFSNI
                     IKYVCKTFPPASQRFVYHKEFLVTRVNTWYCKFTRVDTFTLFRGVYHNNVDCEEFYKA
                     MDDAWHYKKTLAMLNAERTIFKDNAALNFWPKVRDMVIVPLFDASITTGRMSRREVM
                     VKNKDFVYTVLNHIKTYQAKALTYANVLSFVESIRSRVIINGYVARSEWDTDKAILGPL
                     AMTFFLITKLGIHVQDEILLKKFQOKFDRTTNELLIWTSLCDALMGVIPSVKETLVRGGFV
                     KVAEQALEIKVPELYCTFADRLVLQYRKAEEPQSCDLSKPLEEESEKYYNALSELSVLE
                     NLDSFDLEAFKTLCQOKSVDPDMAAKVVVAIMKSELTLPFKKPTEEEISESLKPGEGS
                     SAEHKDVLSLSIQNDAPFPCVKNLVEGSVPAYGMCPKGGGFDKLDVDIADFHLKSVDAVK
                     RGTMMSAVYTGSIEVROMKNYIDYLSASLAATVSNLCKVLRDVHGVDPESOEKSGYWD
                     DMGKLKTLTTCSPNGEPEPNAKVILVDGVPGCGKTKEIIEKVNFSEDLILVPGKEAS
                     VRRGRWMLLKPNAKSHAWGVAEDANHKLVIVLLNWDDGKPVCDETWFRVAVSSDSLIYS
                     KMIIRRNAHAGVIRADKDNVSTVDSFLMHPSRRVFKRLFIDEGLMLHTGCVNFLLLS
                     QCDVAYVYGDTOOIPFICRVANFPYPAHFAKLVADEKEVRRVTLRCPADVTYFLNKY
                     ESETLCDIDLLINVPLDKYRYVGILGAVFTGEWLLVPDFVKGGVTISVIDKRLANSKEC
                     VIGTYRAAAKSKRFQFKLVPNYFVSTVDARKRPWQVHVRIQDLKIEAGWQPLALEVVS
                     VAMVINNVMKGLREKRVAINDPDVEGFEGVVDEFVDSVAAFKAVDNFKRKKKVEEK
                     GVVSKYKYRPEKYAGPDSENLKEENVLQHYKPESVPVFRSGVGRAHSDA"
     CDS
     4877. .5680
                     /function="transport protein"
                     /note="ORF3"
                     /codon_start=1
                     /product="29K"
                     /protein_id="CAA82560.1"
                     /db_xref="GI:488715"
                     /db_xref="SPTREMBL:Q88605"
                     /translation="MSIVSYEPKYSDFLNLSKKEEILPKALTRLKTVSISTKDIISVK
                     ESELCDIDLLINVPLDKYRYVGILGAVFTGEWLLVPDFVKGGVTISVIDKRLANSKEC
                     VIGTYRAAAKSKRFQFKLVPNYFVSTVDARKRPWQVHVRIQDLKIEAGWQPLALEVVS
                     VAMVINNVMKGLREKRVAINDPDVEGFEGVVDEFVDSVAAFKAVDNFKRKKKVEEK
                     PSASEVRNATORVDDATVSIRSOIQLLLSELSGSHGYMNRAEFEALVPWTTAAAT"
     CDS
     5604. .6077
                     /function="coat protein"
                     /note="ORF4"
                     /codon_start=1
                     /product="18K"
                     /protein_id="CAA82561.1"
                     /db_xref="GI:488716"
                     /db_xref="SPTREMBL:Q88606"
                     /translation="MSYNITNPNQYQYFAAVWAEPIPMLNQCISALSOSYQTOAARDT
                     VRQQFSNLLSAVVAPSORPFPEGSRVVNSAVIKPLYEALMKSFDTNRIIEIEEESR
                     VROOFSNLLSAVVAPSORPFPEGSRVVNSAVIKPLYEALMKSFDTNRIIEIEEESR"
```

```
     Query Match 79.3%;     Score 149;     DB 14;     Length 6312;
     Best Local Similarity 100.0%;     Pred. No. 1.8e-74;
     Matches 149;     Conservative 0;     Mismatches 0;     Indels 0;     Gaps 0;

     BASE COUNT     1800 a     1213 c     1590 g     1709 t
     ORIGIN

Qy   27  CGAATTCGTCGATTCGGTTGCACCATTTAAAGCGCGTTGACACCATTTAAAGAAGAAGAAA  5514
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5455  CGAATTCGTCGATTCGGTTGCACCATTTAAAGCGCGTTGACACCATTTAAAGAAGAAGAAA  146

Qy   87  GAAGGTTGAAGAAAAGGGGTCGTAGTAAGTAATACAGACCCGGAGAAGTACCGCCGG  5574
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5515  GAAGGTTGAAGAAAAGGGGTCGTAGTAAGTAATACAGACCCGGAGAAGTACCGCCGG  5574

Qy  147  TCCTGAATTCGTTTAATTTGAAAGAAGAAA  175
```

Right column:

```
Db   5575  TCCTGAATTCGTTTAATTTGAAAGAAGAAA  5603
           ||||||||||||||||||||||||||||||
           ||||||||||||||||||||||||||||||
Db   5575  TCCTGAATTCGTTTAATTTGAAAGAAGAAA  5603

RESULT 3
BRU03387
LOCUS       BRU03387     6311 bp    RNA     linear     VRL 18-FEB-1998
DEFINITION  Turnip vein-clearing virus, complete genome.
ACCESSION   U03387 L22518
VERSION     U03387.1  GI:2894629
KEYWORDS    .
SOURCE      Turnip vein-clearing virus
  ORGANISM  Turnip vein-clearing virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE   1  (bases 68 to 3391)
  AUTHORS   Lartey,R.T., Lane,L.C. and Melcher,U.
  TITLE     Electron microscopic and molecular characterization of turnip
            vein-clearing virus
  JOURNAL   Arch. Virol. 138 (3-4), 287-298 (1994)
  MEDLINE   95091521
   PUBMED   7998035
REFERENCE   2  (bases 1 to 67; 3392 to 6311)
  AUTHORS   Lartey,R.T., Voss,T.C. and Melcher,U.
  TITLE     Completion of a cDNA sequence from a tobamovirus pathogenic to
            crucifers
  JOURNAL   Gene 166, 321-322 (1995)
REFERENCE   3  (bases 1 to 6311)
  AUTHORS   Melcher,U.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-1993) Biochemistry & Molecular Biology, Oklahoma
            State University, NRC 246, Stillwater, OK 74078-0454, USA
REFERENCE   4  (bases 1 to 6311)
  AUTHORS   Melcher,U.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-FEB-1998) Biochemistry & Molecular Biology, Oklahoma
            State University, NRC 246, Stillwater, OK 74078-0454, USA
  REMARK    Sequence update by submitter
            Sequence version replaced gi:619581.
COMMENT     On Feb 18, 1998 this sequence version replaced gi:619581.
FEATURES             Location/Qualifiers
     source          1..6311
                     /organism="Turnip vein-clearing virus"
                     /strain="OSU"
                     /specific_host="Brassica rapa"
                     /db_xref="taxon:29272"
                     /lab_host="Nicotiana tabacum"
                     /note="possibly an isolate of ribgrass mosaic virus"
     CDS             68..473
                     /note="contains RNA polymerase; stop codon at 3389..3391
                     is suppressed under certain conditions and expressed as
                     tyrosine"
                     /codon_start=1
                     /transl_except=(pos:3389..3391,aa:Tyr)
                     /product="RNA replicase read-through component"
                     /protein_id="AAC02783.1"
                     /db_xref="GI:619645"
                     /translation="MAQFQQTIDMQTLQAAAGRNSLVNDLASRRVYDNAVEELNARSR
                     EYQRAAFNNYAENPHFVHCDKPFFQQCELTPAYGTDIYAVALHSIYDIPVEEFGSALLR
                     RPKVHFSKAVSTEQTLIATNAYPEFEISFTHTQSAVHSLAGGLRSLELEYLMMQVPFG
                     SLTYDIGGNFSAHLFKGRDYVHCCMPNLDVRDIARHEGHKEAIHSYVNRLKRQQRPVP
                     KNVKTCFAAFHFHENMLLDCDTVTLDEIGAFQKRAGDNLSFFHNESTLNYTHSFSNI
                     IKYVCKTFPPASQRFVYHKEFLVTRVNTWYCKFTRVDTFTLFRGVYHNNVDCEEFYKA
                     EYQRAAFNNYAENPHFVHCDKPFFQQCELTPAYGTDIYAVALHSIYDIPVEEFGSALLR
                     RPKVHFSKAVSTEQTLIATNAYPEFEISFTHTQSAVHSLAGGLRSLELEYLMMQVPFG
                     SLTYDIGGNFSAHLFKGRDYVHCCMPNLDVRDIARHEGHKEAIHSYVNRLKRQQRPVP
                     MDDAWHYKKTLAMLNAERTIFKDNAALNFWPKVRDMVIVPLFDASITTGRMSRREIM
                     VNKDFVYTVLNHIKTYQAKALTYANVLSFVESIRSRVIINGYVARSEWDTDKAILGPL
                     AMTFFLITKLGIHVQDEILLKKFQOKFDRTTNELLIWTSLCDALMGVIPSVKETLVRGGFV
                     KVAEQALEIKVPELYCTFADRLVLQYRKAEEPQSCDLSKPLEEESEKYYNALSELSVLE
                     NLDSFDLEAFKTLCQOKSVDPDMAAKVVVAIMKSELTLPFKKPTEEEISESLKPGEGS
                     CAEHKEVLSLSIQNDAPFPCVKNLVEGSVPAYGMCPKGGGFDKLDVDIADFHLKSVDAV
                     RGTMMSAVYTGSIKVQOMKNYIDYLSASLAATVSNLCKVLRDVHGVDPESOEKSGYWD
                     VRRGRWMLLKPNAKSHAWGVAEDANHKLVIVLLNWDDGKPVCDETWFRVAVSSDSLIY
                     DMGKLKTLTCSPNGEPEPNAKVILVDGVPGCGKTKEIIEKVNFSEDLILVPGKEAS
                     KMIIRRNAHAGVIRADKDNVSTVDSFLMHPSRRVFKRLFIDEGLMLHTGCVNFLLLS
                     QCDVAYVYGDTOOIPFICRVANFPYPAHFAKLVADEKEVRRVTLRCPADVTYFLNKY
                     DGAVMCISAVERSVKAEVVRGKALNPITLPLEGKILTFTQOADKFELLEKGYKDVI"
```

```
                     HEVQGETKTAIVRLTSTPLEIISSASPHVLVALTRHTTCCKYTVVLDPMVNVISE
                     MEKLSNFLLDMYRVEAGVQYQLQIDAVFRDSNLFVQTFKSGDWRDMQFYDALLPGNS
                     TILNEFDAVTMNLRDISLNVKDCRIDFSKSVQLFKEQPIFLKPKIRTAAEMPRTAGLL
                     ENLVAMIRNRMNAPDLTGTIDIDIEDTASLVVEKFMKSYDYKEFSGTNEMTMRESFSRW
                     LSKQESSTVGQLADFNFVDLPAVDEVKHMIKSDPKQKDLSIQDEYPALQTIVYHSKK
                     INAIFGPMFSELTRMLLERIDSSKFLFYTRKTPAQIEDEFSDLLDSTQAMEIELEDISK
                     YDKSQNEEHCAVEYKIWEKLGIDEWLAEVWKQGHRKTLKDYTAGVKTCLWIQRKSGD
                     VTFPIGNTIIACLSSMIPMDKVIKAAFCGDSLIYIPKGLDLPDIQAGANLMNFE
                     AKLFRKKYGFCGRYVIHHDRGAIYYPDLKLISKLGCKHIRDVVHLEELRESLCDVA
                     SNLNNCAFSQLDEAVAEVHKTAVGGSFRAFCSIIKYLSDKRLFRDLFFV"

     CDS             68..3391
                     /note="contains capping methyltransferase and RNA helicase
                     domains"
                     /codon_start=1
                     /product="RNA replicase component"
                     /protein_id="AAC02782.1"
                     /db_xref="GI:514836"
                     /translation="MAQFQQTIDMQTLQOAAAGNNSLVNDLASRVYDNAVAEELNARSR
                     RPKVHFSKRAVSTEQTLIATNAYPEFEISFTHTQSAVHSLAGGIRSLELEYLMMQVPFG
                     SLIYDIGGNFSAHLFKGRDYVHCCMPNLDVRDIARHEGHKRAIYSYVNRLKRQQRVP
                     EYQRAAFNNYAENPHFYHCDKPFQQCELTTAYGTDTYAVALHSIYDIPVEERGSALLR
                     KNYKTCFAAFHFHIENMLLDCDVTLDEIGATFQKSGDNLSFFHNESTLNYTHSFSNI
                     IKYVCKTFPASQRFVHKEFLVTRVNVDSFLMHPSRVFVKRLFIDEGLMLHTGCVNFLLLS
                     MDDAWHYKKTLAMLNAERTIFKDNAALNFWFKPRVRDMVIVPLFDASITTGRMSRREIM
                     VKRDFVTVLNHIKTVQAKALTANVLSFVESIRSRVIINGVTARSEWDIDKAILGPL
                     AMFFLTFKLGHVQDEIILKKFCDRRTNELIWTSICDALMGVIPSVKETLVRGGFEV
                     KVAEQALEIKVPELYCTFADRLVLQYKRAAEEFQSCDLSKPLEESEKYYNALSELSVLE
                     NLDSFDLEAFKILCQQONVDPDMAAAKVVAIMKSELILPFKKPTEEISESLKPGEGS
                     CAEHKEVLSLQNDAPFPCVKNIVEGSYPAYGMCPKGGGDRHLDVDIADFHLKSVDAVK
                     KGTMMSAVYTGSIKVQOMKNYIDYLSASLAATVSNLCKVLRDVHGVDPESQKSGVWD
                     VRRGRWLLKPNAKSHANGVAEDANHKLVIVLLNWDDGKPVCDETWFWRVAVSSDSLIYS
                     DMGKLRTLTSCSPENGEPEPNAKVLVDGVPGCGKTKEIIEKVNFSEDLILVPGKRAS
                     KMIIRRANQAGVIRADKDNVRTVDSFLMHPSRVFVKRLFIDEGLMLHTGCVNFLLLLS
                     QCDVAYVGDTKQIPFICRVANFPYPAHFAKLVADEKEVRVTLRCPADVIYFLNKKY
                     DGAVMCTSAVERSVKAEVVRGAGALNPIITLPLEGKILTFTQADKFELLEKGYKDVNTV
                     HEVQGETYEKTAIVRLTSTPLEIISSASPHVLVALFRHTTCCKYTVVLDPMVNVISE
                     MEKLSNFLLDMYRVEAGVQ"

     CDS             4876..5679
                     /function="intercellular transport of infection"
                     /codon_start=1
                     /product="movement protein"
                     /protein_id="AAC02784.1"
                     /db_xref="GI:514838"
                     /translation="MSIVSYEPKVSDFLNLSKKEEILPKALIRLKTVSISTRDIISVK
                     ESETLCDIDLLINVPLDKIRYGILGAVFTGEWLVPDFVEVKGGVTISVIDKRLVNSKEC
                     VIGTYRAAAKSKRFQEKLVPNYFVSIVDAKRKPWQVHVRIQDLKIEAGWQPLALEVVS
                     VAMTNNVVMKGLREKVAINDPEVDEGFVDVDEFVDSVAAPKAVDNRFRKRKKKVEER
                     DVVSKYKYRPEKYAGPDSFNLKEENVLQHYKPESVPVLRSGVGRAHTNA"

     CDS             5603..6076
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="AAC02785.1"
                     /db_xref="GI:514839"
                     /translation="MSYNITNPNQYQYFAAVWAEPIPMLNQCMSALSQSYQTQAARDT
                     PSASEVANATQRVDDATVAIRSQIQLLSELSNGHGYMNRAEFEALLPWTTAPAT"

BASE COUNT     1810 a   1195 c   1586 g   1720 t
ORIGIN
```

Query Match   22.3%; Score 42; DB 14; Length 6311;
Best Local Similarity 100.0%; Pred. No. 8.8e-13;
Matches   42; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     27  CGAATTCGTTCGATTCGGTTGCAGCATTTAAAGCGGTTGACAA  68
           |||||||||||||||||||||||||||||||||||||||||||
Db   5454  CGAATTCGTTCGATTCGGTTGCAGCATTTAAAGCGGTTGACAA  5495
```

RESULT 4

```
MDVFRA
LOCUS       MDVFRA                  1568 bp    mRNA       linear   VRL 13-APR-1995
DEFINITION  Gallid herpesvirus mRNA fragment.
ACCESSION   L41561
VERSION     L41561.1  GI:767682
KEYWORDS    
SOURCE      Gallid herpesvirus (strain RPL-1) cDNA to mRNA.
  ORGANISM  Gallid herpesvirus
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae.
REFERENCE   1  (bases 1 to 1568)
  AUTHORS   Li,D.; Greenall,L.; Smith,G.D.; O'Sullivan,G. and Ross,N.
  TITLE     Marek's disease virus irs-specific cDNA clones are derived from a
            single transcription unit complementary to the icp4 gene
  JOURNAL   Unpublished (1995)
FEATURES             Location/Qualifiers
     source          1..1568
                     /organism="Gallid herpesvirus"
                     /strain="RPL-1"
                     /db_xref="taxon:39566"
                     /haplotype="US"
                     /cell_line="RPL-1"
                     /cell_type="lymphoblast"
BASE COUNT      389 a    371 c    474 g    334 t
ORIGIN
```

Query Match   20.2%; Score 38; DB 14; Length 1568;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches   38; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1  GTCGGACGGTATCGATAAGCTTGATATCGAATTCGTCGCA  38
           ||||||||||||||||||||||||||||||||||||||
Db      4  GTCGGACGGTATCGATAAGCTTGATATCGAATTCGTCGA  41
```

RESULT 5

```
AR067739
LOCUS       AR067739                2126 bp    DNA        linear   PAT 29-SEP-1999
DEFINITION  Sequence 1 from patent US 5851798.
ACCESSION   AR067739
VERSION     AR067739.1  GI:5998961
KEYWORDS    
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2126)
  AUTHORS   Shabon,U. and Bergsma,D.
  TITLE     Nucleic acid encoding human GPR14 receptor
  JOURNAL   Patent: US 5851798-A 1 22-DEC-1998.
FEATURES             Location/Qualifiers
     source          1..2126
                     /organism="unknown"
BASE COUNT      317 a    763 c    669 g    377 t
ORIGIN
```

Query Match   18.6%; Score 35; DB 6; Length 2126;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches   35; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1  GTCGGACGGTATCGATAAGCTTGATATCGAATTCGT  35
           ||||||||||||||||||||||||||||||||||||
Db     30  GTCGGACGGTATCGATAAGCTTGATATCGAATTCGT  64
```

RESULT 6

```
AR096004
LOCUS       AR096004                2126 bp    DNA        linear   PAT 08-SEP-2000
DEFINITION  Sequence 1 from patent US 6005074.
ACCESSION   AR096004
VERSION     AR096004.1  GI:10024409
KEYWORDS    
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2126)
  AUTHORS   Shabon,U. and Bergsma,D.
  TITLE     Cloning of human GPR14 receptor
  JOURNAL   Patent: US 6005074-A 1 21-DEC-1999;
```

```
Qy      1  GTCGGACGGTATCGATAAGCTTGATATCGAATTCGT  35
           ||||||||||||||||||||||||||||||||||||
Db     30  GTCGGACGGTATCGATAAGCTTGATATCGAATTCGT  64
```

RESULT 7
```
LOCUS           AR116595                2126 bp    DNA         linear   PAT 16-MAY-2001
DEFINITION      Sequence 1 from patent US 6133420.
ACCESSION       AR116595
VERSION         AR116595.1  GI:14096917
KEYWORDS        .
SOURCE          Unknown.
  ORGANISM      Unknown.
                Unclassified.
REFERENCE       1  (bases 1 to 2126)
  AUTHORS       Ames,R.S. Jr., Sarau,H.M., Foley,J.J., Shabon,U., Bergsma,D. and
                Chambers,J.K.
  TITLE         GPR14 polypeptides
  JOURNAL       Patent: US 6133420-A 1 17-OCT-2000;
FEATURES             Location/Qualifiers
     source          1..2126
                     /organism="unknown"
BASE COUNT       317 a    763 c    669 g    377 t
ORIGIN
Qy          1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
              |||||||||||||||||||||||||||||||||||
Db         30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

Query Match   18.6%; Score 35; DB 6; Length 2126;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 35; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 8
```
LOCUS           AR121154                2126 bp    DNA         linear   PAT 16-MAY-2001
DEFINITION      Sequence 1 from patent US 6159700.
ACCESSION       AR121154
VERSION         AR121154.1  GI:14104730
KEYWORDS        .
SOURCE          Unknown.
  ORGANISM      Unknown.
                Unclassified.
REFERENCE       1  (bases 1 to 2126)
  AUTHORS       Aiyar,N.V., Ames,R.S., Arnold,A.Romanic., Al-Barazanji,K.,
                Bergsma,D.J., Chambers,J., Douglas,S.A., Foley,J.J., Gout,B.,
                Khandoudi,N., Sarau,H.M., Shabon,U. and Willette,R.N.
  TITLE         Method of finding agonist and antagonist to human and rat GPR14
  JOURNAL       Patent: US 6159700-A 112-DEC-2000;
FEATURES             Location/Qualifiers
     source          1..2126
                     /organism="unknown"
BASE COUNT       317 a    763 c    669 g    377 t
ORIGIN
Qy          1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
              |||||||||||||||||||||||||||||||||||
Db         30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

Query Match   18.6%; Score 35; DB 6; Length 2126;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 35; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Db         30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

RESULT 9
```
LOCUS           AR153837                7560 bp    DNA         linear   PAT 08-AUG-2001
DEFINITION      Sequence 4 from patent US 6235975.
ACCESSION       AR153837
VERSION         AR153837.1  GI:15121369
KEYWORDS        .
SOURCE          Unknown.
  ORGANISM      Unknown.
                Unclassified.
REFERENCE       1  (bases 1 to 7560)
  AUTHORS       Harada,J.J., Lotan,T., Ohto,M.-a., Goldberg,R.B. and Fischer,R.L.
  TITLE         Leafy cotyledon1 genes and methods of modulating embryo development
                in transgenic plants
  JOURNAL       Patent: US 6235975-A 4 22-MAY-2001;
FEATURES             Location/Qualifiers
     source          1..7560
                     /organism="unknown"
BASE COUNT      2378 a   1326 c   2573 t     19 others
ORIGIN
Qy          1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
              |||||||||||||||||||||||||||||||||||
Db         53 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 87
```

Query Match   18.6%; Score 35; DB 6; Length 7560;
Best Local Similarity 100.0%; Pred. No. 9.8e-09;
Matches 35; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 10
```
LOCUS           AB027145                 110 bp    mRNA        linear   ROD 06-JUL-2000
DEFINITION      Rattus norvegicus mRNA for CDCrel-1F, partial cds.
ACCESSION       AB027145
VERSION         AB027145.1  GI:8953678
KEYWORDS        CDCrel-1/PNUTL1 fetus form.
SOURCE          Rattus norvegicus (strain:wistar) neonatal male neocortex cDNA to
                mRNA.
  ORGANISM      Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE       1  (sites)
  AUTHORS       Toda,S., Kajii,Y., Sato,M. and Nishikawa,T.
  TITLE         Reciprocal expression of infant- and adult-preferring transcripts
                of CDCrel-1 septin gene in the rat neocortex
  JOURNAL       Biochem. Biophys. Res. Commun. 273 (2), 723-728 (2000)
  MEDLINE       20334336
REFERENCE       2  (bases 1 to 110)
  AUTHORS       Toda,S., Kajii,Y. and Nishikawa,T.
  TITLE         Direct Submission
  JOURNAL       Submitted (11-MAY-1999) Shigenobu Toda, National Institute of
                Neuroscience,NCNP, Department of Mental Disorder Research; 4-1-1
                Ogawahigashi, Kodaira, Tokyo 187-8502; Japan
                (E-mail:toda@psy.med.tohoku.ac.jp, Tel:81-423-46-1714,
                Fax:81-423-46-1744)
FEATURES             Location/Qualifiers
     source          1..110
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /tissue_type="neocortex"
                     /dev_stage="neonatal"
     gene            1..110
                     /gene="CDCrel-1/PNUTL1f"
     CDS             57..>110
                     /gene="CDCrel-1/PNUTL1f"
                     /codon_start=1
```

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
              |||||||||||||||||||||||||||||||||||
Db   53 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 87

```
                /product="CDcre1-1F"
                /protein_id="BAA98052.1"
                /db_xref="GI:8953679"
                /translation="MSTGLRYKSKLATPEDKQ"
BASE COUNT     30 a    31 c    31 g    18 t
ORIGIN

Query Match     18.1%;  Score 34;  DB 10;  Length 110;
Best Local Similarity   100.0%;  Pred. No. 2.4e-08;
Matches  34;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 GTCGACGGGTATCGATAAGCTTGATATCGAATTCG 34
        |||||||||||||||||||||||||||||||||||
Db  20 GTCGACGGGTATCGATAAGCTTGATATCGAATTCG 53
```

```
RESULT 11
CSPAG5R/c
LOCUS       CSPAG5R        338 bp     DNA      linear   BCT 04-JAN-1995
DEFINITION  Calothrix D253 genomic DNA (clone AG5R).
ACCESSION   Z47159
VERSION     Z47159.1  GI:619119
KEYWORDS    .
SOURCE      Calothrix sp.
  ORGANISM  Calothrix sp.
            Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Calothrix.
REFERENCE   1 (bases 1 to 338)
  AUTHORS   Robinson,N.J., Robinson,P.J. and Gupta,A.
  TITLE     Genomic DNA sequence obtained in the process of generating a random
            sequence database greater than 10 kb from single sequence runs from
            random clones of a Calothrix D253 library
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 338)
  AUTHORS   Robinson,N.J., Robinson,P.J., Gupta,A., Bleasby,A.J., Whitton,B.A.
            and Morby,A.P.
  TITLE     Singular over-representation of HIP1 in DNA from many cyanobacteria
  JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 338)
  AUTHORS   Robinson,N.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1994) Nigel J. Robinson, Biochemistry and
            Genetics, The Medical School, University of Newcastle, Framlington
            Place, Newcastle upon Tyne, NE2 4HH, UK
REFERENCE   4 (bases 1 to 338)
  AUTHORS   Robinson,N.J., Robinson,P.J., Gupta,A., Bleasby,A.J., Whitton,B.A.
            and Morby,A.P.
  TITLE     Singular over-representation of an octameric palindrome, HIP1, in
            DNA from many cyanobacteria
  JOURNAL   Nucleic Acids Res. 23 (5), 729-735 (1995)
  MEDLINE   95523777
   PUBMED   7708486
FEATURES             Location/Qualifiers
     source          1..338
                     /organism="Calothrix sp."
                     /strain="D253"
                     /db_xref="taxon:1187"
                     /clone="AG5R"
                     /clone_lib="LzAPcalothrix-LIB"
BASE COUNT     83 a    87 c    77 g    91 t
ORIGIN

Query Match     18.1%;  Score 34;  DB 1;  Length 338;
Best Local Similarity   100.0%;  Pred. No. 2.7e-08;
Matches  34;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 GTCGACGGGTATCGATAAGCTTGATATCGAATTCG 34
         |||||||||||||||||||||||||||||||||||
Db  138 GTCGACGGGTATCGATAAGCTTGATATCGAATTCG 105
```

```
RESULT 12
G16087/c
LOCUS       G16087         369 bp     DNA      linear   STS 13-FEB-1996
```

```
DEFINITION  chicken STS ADL320, sequence tagged site.
ACCESSION   G16087
VERSION     G16087.1  GI:1184628
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Gallus gallus
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 369)
  AUTHORS   Cheng,H.H.
  TITLE     Unpublished (1994)
  JOURNAL   Unpublished
COMMENT     Synonyms: B664
            Contact: Hans H. Cheng
            Avian Disease and Oncology Laboratory
            USDA-ARS
            3606 E. Mount Hope Rd, East Lansing, MI 48823, USA
            Tel: 517376758
            Fax: 517376776
            Email: hcheng@pilot.msu.edu
            Primer A: AGGGGTTATTGCCTGCTCTGC
            Primer B: GTCCTCAGTGTGCCAAATGC
            STS size: 114
            PCR Profile:
            Denaturation:    94 degrees C for 1 minute
            Annealing:       51 degrees C for 1 minute
            Polymerization:  72 degrees C for 1 minute
            PCR Cycles:      30
            Thermal Cycler:  MJ Research
            Protocol:
            Template:        50-100  ng
            Primer:          0.1  uM
            dNTPs:           200  uM
            Taq polymerase:  0.4 units/ul
            Total Vol:       25  ul
            Buffer:
            MgCl2:           1.5  mM
            KCl:             50   mM
            Tris-HCl:        10   mM
            pH:              9.0
            Triton X-100:    0.1  %.
FEATURES             Location/Qualifiers
     source          1..369
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
     STS             35..148
     primer_bind     35..54
     primer_bind     complement(129..148)
BASE COUNT     96 a    54 c    112 g    107 t
ORIGIN

Query Match     18.1%;  Score 34;  DB 11;  Length 369;
Best Local Similarity   100.0%;  Pred. No. 2.7e-08;
Matches  34;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 GTCGACGGGTATCGATAAGCTTGATATCGAATTCG 34
         |||||||||||||||||||||||||||||||||||
Db  323 GTCGACGGGTATCGATAAGCTTGATATCGAATTCG 290
```

```
RESULT 13
G01691/c
LOCUS       G01691         373 bp     DNA      linear   STS 07-DEC-1994
DEFINITION  chicken STS ADL271, sequence tagged site.
ACCESSION   G01691
VERSION     G01691.1  GI:595215
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Gallus gallus
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
```

```
REFERENCE
  1  (bases 1 to 373)
  AUTHORS   Cheng,H.H.
  JOURNAL   Unpublished (1994)
          Phasianinae; Gallus.
          Synonyms: B367
COMMENT   Contact: Hans H. Cheng
          Avian Disease and Oncology Laboratory
          USDA-ARS
          3606 E. Mount Hope Rd, East Lansing, MI 48823, USA
          Tel: 517376758
          Fax: 517376776
          Email: hcheng@pilot.msu.edu
          Primer A: TCATCAGAGAGCCCAAGCACA
          Primer B: ATGAATGAAACCCATCTAAC
          STS size: 136
          PCR Profile:
          Denaturation:    94 degrees C for 1 minute
          Annealing:       48 degrees C for 1 minute
          Polymerization:  72 degrees C for 1 minute
          PCR Cycles:      30
          Thermal Cycler:  MJ Research
          Protocol:
          Template:        50-100  ng
          Primer:          0.1  uM
          Primer:          0.1  uM
          dNTPs:           200  uM
          Taq polymerase:  0.4  units/ul
          Total Vol:       25  ul
          Buffer:
          MgCl2:           1.5  mM
          KCl:             50  mM
          Tris-HCl:        10  mM
          pH:              9.0
          Triton X-100:    0.1 %.
FEATURES          Location/Qualifiers
     source         1..373
                    /organism="Gallus gallus"
                    /db_xref="taxon:9031"
     STS            1..138
     primer_bind    3..22
     primer_bind    complement(119..138)
BASE COUNT    98 a    81 c    81 g    113 t
ORIGIN

Qy          1 GTCGACGGGTATCGATAAGCTTGAATCGAATTCG  34
              ||||||||||||||||||||||||||||||||||
Db        257 GTCGACGGGTATCGATAAGCTTGAATCGAATTCG 224

Query Match             18.1%;  Score 34;  DB 11;  Length 373;
  Best Local Similarity  100.0%;  Pred. No. 2.7e-08;
  Matches 34;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;


RESULT 14
LOCUS       CSPAG13R/c
            CSPAG13R    397 bp    DNA    linear    BCT 04-JAN-1995
DEFINITION  Calothrix D253 genomic DNA (clone AG13R).
ACCESSION   247173
VERSION     247173.1  GI:619086
KEYWORDS    .
SOURCE      Calothrix sp.
  ORGANISM  Calothrix sp.
            Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Calothrix.
REFERENCE   1  (bases 1 to 397)
  AUTHORS   Robinson,N.J., Robinson,P.J. and Gupta,A.
  TITLE     Genomic DNA sequence obtained in the process of generating a
            random sequence database greater than 10 kb from single sequence
            runs from random clones of a Calothrix D253 library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 397)
  AUTHORS   Robinson,N.J., Robinson,P.J., Gupta,A., Bleasby,A.J., Whitton,B.A.
            and Morby,A.P.
  TITLE     Singular over-representation of HIP1 in DNA from many cyanobacteria
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 397)
  AUTHORS   Robinson,N.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1994) Nigel J. Robinson, Biochemistry and
            Genetics, The Medical School, University of Newcastle, Framlington
            Place, Newcastle upon Tyne, NE2 4HH, UK
REFERENCE   4  (bases 1 to 397)
  AUTHORS   Robinson,N.J., Robinson,P.J., Gupta,A., Bleasby,A.J., Whitton,B.A.
            and Morby,A.P.
  TITLE     Singular over-representation of an octameric palindrome, HIP1, in
            DNA from many cyanobacteria
  JOURNAL   Nucleic Acids Res. 23 (5), 729-735 (1995)
  MEDLINE   95223777
   PUBMED   7708486
FEATURES          Location/Qualifiers
     source         1..397
                    /organism="Calothrix sp."
                    /strain="D253"
                    /db_xref="taxon:1187"
                    /clone="AG13R"
                    /clone_lib="L2APCalothrix-LIB"
BASE COUNT   105 a    91 c    88 g    113 t
ORIGIN

Qy          1 GTCGACGGGTATCGATAAGCTTGAATCGAATTCG  34
              ||||||||||||||||||||||||||||||||||
Db         90 GTCGACGGGTATCGATAAGCTTGAATCGAATTCG  57

Query Match             18.1%;  Score 34;  DB 1;  Length 397;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-08;
  Matches 34;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;


RESULT 15
LOCUS       HRSTCRAG
            HRSTCRAG    487 bp    DNA    linear    MAM 01-MAY-1995
DEFINITION  Equus caballus (clone T131) T-cell receptor DNA, V-region.
ACCESSION   L27850
VERSION     L27850.1  GI:790760
KEYWORDS    T-cell receptor; V-region.
SOURCE      Equus caballus (clone: T131) DNA.
  ORGANISM  Equus caballus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   Schrenzel,M.D., Watson,J.L. and Ferrick,D.A.
  TITLE     Characterization of horse (Equus caballus) T-cell receptor beta
            chain genes
  JOURNAL   Immunogenetics 40 (2), 135-144 (1994)
  MEDLINE   94292284
   PUBMED   7913080
FEATURES          Location/Qualifiers
     source         1..487
                    /organism="Equus caballus"
                    /db_xref="taxon:9796"
                    /clone="T131"
     CDS            <1..>430
                    /note="This CDS feature is included to show the
                    translation of the corresponding V_region. Presently
                    translation qualifiers on V_region features are illegal."
                    /codon_start=2
                    /product="T cell receptor"
                    /protein_id="AAA65661.1"
                    /db_xref="GI:790761"
                    /translation="RSTVSILSINSARGHISLERGVLYSKPLKMLMFLLLGPTGS
                    GLGVLVSQHPSRAICKNGSSVEMECRTVDLQAQTCWYQQWANQGITPVATSNQGLSS
                    TYEQGPFKVKFPISHPNLTFSTLTICAHPANSGLYFSCAR"
     V_region       <1..430
                    /product="T cell receptor"
     J_segment      431..481
                    /note="DJ"
BASE COUNT   114 a   132 c   123 g   118 t
```

ORIGIN

Query Match   18.1%;   Score 34;   DB 4;   Length 487;
Best Local Similarity   100.0%;   Pred. No. 2.8e-08;
Matches   34;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGGAATTCG   34
       ||||||||||||||||||||||||||||||||||
Db   7 GTCGACGGTATCGATAAGCTTGATAATCGAATTCG   40

Search completed: February 19, 2003, 07:57:01
Job time : 2608 secs

OM nucleic - nucleic search, using sw model

Run on: February 19, 2003, 06:07:16 ; Search time 229 Seconds
                                                (without alignments)
                                                1848.802 Million cell updates/sec

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.3

Title: US-09-758-962-1
Perfect score: 188
Sequence: 1 gtcgacgtatcgataagct..........gaagaaagagctcaccatgg 188

Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size :   0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :   N_Geneseq_101002:*
          1: /SIDS2/gcgdata/geneseqn-embl/NA1980.DAT:*
          2: /SIDS2/gcgdata/geneseqn-embl/NA1981.DAT:*
          3: /SIDS2/gcgdata/geneseqn-embl/NA1982.DAT:*
          4: /SIDS2/gcgdata/geneseqn-embl/NA1983.DAT:*
          5: /SIDS2/gcgdata/geneseqn-embl/NA1984.DAT:*
          6: /SIDS2/gcgdata/geneseqn-embl/NA1985.DAT:*
          7: /SIDS2/gcgdata/geneseqn-embl/NA1986.DAT:*
          8: /SIDS2/gcgdata/geneseqn-embl/NA1987.DAT:*
          9: /SIDS2/gcgdata/geneseqn-embl/NA1988.DAT:*
         10: /SIDS2/gcgdata/geneseqn-embl/NA1989.DAT:*
         11: /SIDS2/gcgdata/geneseqn-embl/NA1990.DAT:*
         12: /SIDS2/gcgdata/geneseqn-embl/NA1991.DAT:*
         13: /SIDS2/gcgdata/geneseqn-embl/NA1992.DAT:*
         14: /SIDS2/gcgdata/geneseqn-embl/NA1993.DAT:*
         15: /SIDS2/gcgdata/geneseqn-embl/NA1994.DAT:*
         16: /SIDS2/gcgdata/geneseqn-embl/NA1995.DAT:*
         17: /SIDS2/gcgdata/geneseqn-embl/NA1996.DAT:*
         18: /SIDS2/gcgdata/geneseqn-embl/NA1997.DAT:*
         19: /SIDS2/gcgdata/geneseqn-embl/NA1998.DAT:*
         20: /SIDS2/gcgdata/geneseqn-embl/NA1999.DAT:*
         21: /SIDS2/gcgdata/geneseqn-embl/NA2000.DAT:*
         22: /SIDS2/gcgdata/geneseqn-embl/NA2001A.DAT:*
         23: /SIDS2/gcgdata/geneseqn-embl/NA2001B.DAT:*
         24: /SIDS2/gcgdata/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match Length	DB	ID	Description
1	100.0	188	188	24	ABQ76058	Clone TXS.GFP-IRES
2	96.3	181	181	24	ABQ76062	Clone TXS.GFP-IRES
c	3	79.3	149	187	24	ABQ76060	Clone TXS.GFP-SERI
4	79.3	149	251	24	ABQ76059	Clone TXS.GFP-HIRE
5	23.4	44	180	24	ABQ76063	Clone TXS.GFP-IRES
6	22.9	43	181	24	ABQ76064	Clone SC196 contai
7	18.6	35	122	24	ABQ76061	Clone SC197 contai
c	8	18.6	35	444	23	ABV36222	Human prostate exp
c	9	18.6	35	444	23	ABV45268	Human prostate exp

c	10	35	18.6	514	23	ABV49873	Human prostate exp
11	35	18.6	1353	24	ABK52453	Accessory factor T
12	35	18.6	2126	19	AAV41999	Human GPR14 (G-pro
13	35	18.6	2126	20	AAZ11174	Human GPR14 coding
14	35	18.6	3236	24	AAK99298	cDNA of the human
15	35	18.6	3905	24	ABN83030	Human nervous syst
16	35	18.6	7560	19	AAV61024	Human X2 theca rep
17	35	18.6	7560	21	AAZ48840	Arabidopsis leafy-
18	35	18.6	7560	21	AAD16916	7.4 kb EcpRI fragm
19	34	18.1	99	23	ABL53457	Plasmid NBE-T frag
20	34	18.1	353	22	ABA11976	Human nervous syst
21	34	18.1	361	21	AAC59649	Human secreted pro
22	34	18.1	362	20	AAF16468	Human prostate can
23	34	18.1	420	20	AAX37517	Human secreted pro
24	34	18.1	421	22	ABA13348	Human nervous syst
25	34	18.1	470	22	AAX37520	Human secreted pro
26	34	18.1	491	22	ABA13327	Human nervous syst
27	34	18.1	501	22	ABA14279	Human secreted pro
28	34	18.1	502	21	AAC59628	Human secreted pro
29	34	18.1	508	22	ABA13368	Human nervous syst
30	34	18.1	523	22	ABA13305	Human nervous syst
31	34	18.1	528	22	ABA14286	Human secreted pro
32	34	18.1	585	22	ABA13335	Human nervous syst
33	34	18.1	587	22	AAX37393	Human secreted pro
34	34	18.1	591	24	ABN90021	Mouse clone IMX3_6
35	34	18.1	602	22	ABA13360	Human nervous syst
36	34	18.1	607	22	ABA14282	Human nervous syst
37	34	18.1	610	22	ABA13372	Human nervous syst
38	34	18.1	616	20	AAX22262	Human secreted pro
39	34	18.1	626	22	AAS41400	cDNA encoding nove
40	34	18.1	642	22	ABA13400	Human nervous syst
41	34	18.1	655	22	ABA14170	Human nervous syst
42	34	18.1	712	22	AAD05161	Human secreted pro
43	34	18.1	723	22	ABA14285	Human nervous syst
c	44	34	18.1	740	22	AAF45087	Human secreted pro
45	34	18.1	741	22	AAH55734	Human breast tumou

ALIGNMENTS

RESULT 1
ABQ76058
ID   ABQ76058 standard; DNA; 188 BP.
XX
AC   ABQ76058;
XX
DT   30-SEP-2002 (first entry)
XX
DE   Clone TXS.GFP-IRES8-CP containing IRES DNA sequence.
XX
KW   IRES; internal ribosome entry site; viral protein; bicistronic;
KW   virus infection; coat protein; protein replacement therapy; crop;
KW   nutritional value; seed oil content; ds.
XX
OS   Synthetic.
XX
PN   WO20025719-A2.
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2002; 2002WO-US01123.
XX
PR   09-JAN-2001; 2001US-0758962.
XX
PA   (BIOS-) BIOSOURCE GENETICS CORP.
XX
PI   Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX
DR   WPI; 2002-557829/59.
XX
PT   New polypeptides contained in plant virus expression vectors, as gene
PT   expression tools, in protein replacement therapy or for intervening in

PT   a metabolic pathway to improve the nutritional value of a crop or alter
PT   the oil content of seeds
XX
PS   Claim 4; Fig 3; 33pp; English.
XX
CC   This invention describes a novel isolated polynucleotide comprising an
CC   internal ribosome entry site (IRES) nucleotide sequence, an open reading
CC   frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,
CC   where the IRES is located between ORF1 and ORF2. The novel
CC   polynucleotides of the invention are used in the construction of a
CC   recombinant potato virus X-based viral vector containing a nucleic acid
CC   construct comprising a bicistronic message with an intervening IRES. The
CC   constructs are used in a method for regulating the rate at which a virus
CC   infection spreads in a host. Regulation is achieved by placing the
CC   nucleic acid construct comprising an internal ribosomes entry site
CC   upstream of a coat protein gene, where the IRES is chosen by the rate of
CC   infection of the viral vector on a host in the presence of that IRES. The
CC   polynucleotide and vectors of the invention are useful for directing
CC   rapid and high-level expression of foreign genes in mature,
CC   differentiated, plant tissue. These are particularly useful in protein
CC   replacement therapy, or for intervening in a metabolic pathway to improve
CC   the nutritional value of a crop or alter the oil content of the seed.
CC   This sequence represents a fragment of the clone TXS.GFP.IRESs-CP
CC   containing an IRES region described in the method of the invention.
XX
SQ   Sequence 188 BP; 66 A; 26 C; 51 G; 45 T; 0 other;

Query Match   100.0%; Score 188; DB 24; Length 188;
Best Local Similarity   100.0%; Pred. No. 1.3e-81;
Matches   188; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTCGATTCGGTTGCAGCATTTAAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTCGATTCGGTTGCAGCATTTAAAGCG 60

Qy   61 GTTGACAACTTTAAAGAAGNAAGGAAAAAGNAGGGTTGAAGAAAAAGNGGTGTTAGTAAGTAAGTAT 120
              |||||||||||||||||| |||||||||| ||||||||||||||| ||||||||||||||||||
Db   61 GTTGACAACTTTAAAGAAGGAAGGAAAAAGGAGGGTTGAAGAAAAAGGGGTGTTAGTAAGTAAGTAT 120

Qy   121 AAGTACAGACCGGAGAAGTACGCCGGGTCCTGATTCGTTTAATTTGAAAGAAGAAAAGAGCT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAGTACAGACCGGAGAAGTACGCCGGGTCCTGATTCGTTTAATTTGAAAGAAGAAAAGAGCT 180

RESULT 2
ABQ76062
ID   ABQ76062 standard; DNA; 181 BP.
XX
AC   ABQ76062;
XX
DT   30-SEP-2002 (first entry)
XX
DE   Clone TXS.GFP-IRESs-CP containing IRESCp DNA sequence.
XX
KW   IRES; internal ribosome entry site; viral protein; bicistronic;
KW   virus infection; coat protein; protein replacement therapy; crop;
KW   nutritional value; seed oil content; ds.
XX
OS   Synthetic.
XX
PN   WO200255719-A2.
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2002; 2002WO-US01123.
XX
PR   09-JAN-2001; 2001US-0758962.
XX
PA   (BIOS-) BIOSOURCE GENETICS CORP.

Qy   181 CACCATGG 188
              ||||||||
Db   181 CACCATGG 188

XX
PI   Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX
DR   WPI; 2002-557829/59.
XX
PT   New polypeptides contained in plant virus expression vectors, as gene
PT   expression tools, in protein replacement therapy or for intervening in
PT   a metabolic pathway to improve the nutritional value of a crop or alter
PT   the oil content of seeds
XX
PS   Claim 4; Fig 4; 33pp; English.
XX
CC   This invention describes a novel isolated polynucleotide comprising an
CC   internal ribosome entry site (IRES) nucleotide sequence, an open reading
CC   frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,
CC   where the IRES is located between ORF1 and ORF2. The novel
CC   polynucleotides of the invention are used in the construction of a
CC   recombinant potato virus X-based viral vector containing a nucleic acid
CC   construct comprising a bicistronic message with an intervening IRES. The
CC   constructs are used in a method for regulating the rate at which a virus
CC   infection spreads in a host. Regulation is achieved by placing the
CC   nucleic acid construct comprising an internal ribosomes entry site
CC   upstream of a coat protein gene, where the IRES is chosen by the rate of
CC   infection of the viral vector on a host in the presence of that IRES. The
CC   polynucleotide and vectors of the invention are useful for directing
CC   rapid and high-level expression of foreign genes in mature,
CC   differentiated, plant tissue. These are particularly useful in protein
CC   replacement therapy, or for intervening in a metabolic pathway to improve
CC   the nutritional value of a crop or alter the oil content of the seed.
CC   This sequence represents a fragment of the clone TXS.GFP.IRESs-CP
CC   containing an IREScp region described in the method of the invention.
XX
SQ   Sequence 181 BP; 64 A; 24 C; 49 G; 44 T; 0 other;

Query Match   96.3%; Score 181; DB 24; Length 181;
Best Local Similarity   100.0%; Pred. No. 3.2e-78;
Matches   181; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTCGATTCGGTTGCAGCATTTAAAGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTCGATTCGGTTGCAGCATTTAAAGCG 60

Qy   61 GTTGACAACTTTAAAGAAGGAAGGAAAAAGGAGGGTTGAAGAAAAAGGGGTGTTAGTAAGTAAGTAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTTGACAACTTTAAAGAAGGAAGGAAAAAGGAGGGTTGAAGAAAAAGGGGTGTTAGTAAGTAAGTAT 120

Qy   121 AAGTACAGACCGGAGAAGTACGCCGGGTCCTGATTCGTTTAATTTGAAAGAAGAAAAGAGCT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAGTACAGACCGGAGAAGTACGCCGGGTCCTGATTCGTTTAATTTGAAAGAAGAAAAGAGCT 180

Qy   181 C 181
              |
Db   181 C 181

RESULT 3
ABQ76060/c
ID   ABQ76060 standard; DNA; 187 BP.
XX
AC   ABQ76060;
XX
DT   30-SEP-2002 (first entry)
XX
DE   Clone TXS.GFP-SERI-CP containing IRES DNA sequence.
XX
KW   IRES; internal ribosome entry site; viral protein; bicistronic;
KW   virus infection; coat protein; protein replacement therapy; crop;
KW   nutritional value; seed oil content; ds.
XX
OS   Synthetic.
XX
PN   WO200255719-A2.

PD   18-JUL-2002.
XX
PF   09-JAN-2002; 2002WO-US01123.
XX
PR   09-JAN-2001; 2001US-0758962.
XX
PA   (BIOS-) BIOSOURCE GENETICS CORP.
XX
PI   Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX
DR   WPI; 2002-557829/59.
XX
PS   Claim 4; Fig 3; 33pp; English.
XX
PT   New polypeptides contained in plant virus expression vectors, as gene
PT   expression tools, in protein replacement therapy or for intervening in
PT   a metabolic pathway to improve the nutritional value of a crop or alter
PT   the oil content of seeds -
XX
CC   This invention describes a novel isolated polynucleotide comprising an
CC   internal ribosome entry site (IRES) nucleotide sequence, an open reading
CC   frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,
CC   where the IRES is located between ORF1 and ORF2. The novel
CC   polynucleotides of the invention are used in the construction of a
CC   recombinant potato virus X-based viral vector containing a nucleic acid
CC   construct comprising a bicistronic message with an intervening IRES. The
CC   constructs are used in a method for regulating the rate at which a virus
CC   infection spreads in a host. Regulation is achieved by placing the
CC   nucleic acid construct comprising an internal ribosome entry site
CC   upstream of a coat protein gene, where the IRES is chosen by the rate of
CC   infection of the viral vector on a host in the presence of that IRES. The
CC   polynucleotide and vectors of the invention are useful for directing
CC   rapid and high-level expression of foreign genes in mature,
CC   differentiated, plant tissue. These are particularly useful in protein
CC   replacement therapy, or for intervening in a metabolic pathway to improve
CC   the nutritional value of a crop or alter the oil content of the seed.
CC   This sequence represents a fragment of the clone TXS.GFP.SERI-CP
CC   containing an IRES region described in the method of the invention.
XX
SQ   Sequence 187 BP; 45 A; 47 C; 27 G; 68 T; 0 other;

Query Match   79.3%; Score 149; DB 24; Length 187;
Best Local Similarity   100.0%;
Matches 149; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   29 AATTCGTCGATTCGGTTGCAGCGGTTGCACACTTTAAAGCGGTTGACACACTTTAAAGAGGAAGGAAGAAAGA 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 AATTCGTCGATTCGGTTGCAGCGGTTGCACACTTTAAAGCGGTTGACAACTTTAAAGAGGAAGGAAGAAAGA 121

Qy   89 AGGTTGAAGAAAAGGGTGTAGTAAGTAAGTATACAGACCGGAGAAGTACGCCGGTC 148
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120 AGGTTGAAGAAAAGGGTGTACTAAGTAAGTATAAGTACAGACCGGAGAAGTACGCCGGTC 61

Qy   149 CTGATTCGTTTAATTGAAAGAAGAAAGA 177
              ||||||||||||||||||||||||||||
Db   60 CTGATTCGTTTAATTGAAAGAAGAAAGA 32

RESULT 4
ID   ABQ76059 standard; DNA; 251 BP.
XX
AC   ABQ76059;
XX
DT   30-SEP-2002 (first entry)
XX
DE   Clone TXS.GFP-HIRES-CP containing IRES DNA sequence.
XX
KW   IRES; internal ribosome entry site; viral protein; bicistronic;
KW   virus infection; coat protein; protein replacement therapy; crop;
KW   nutritional value; seed oil content; ds.
XX
OS   Synthetic.

XX
PN   WO200255719-A2.
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2002; 2002WO-US01123.
XX
PA   (BIOS-) BIOSOURCE GENETICS CORP.
XX
PI   Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX
DR   WPI; 2002-557829/59.
XX
PS   Claim 4; Fig 3; 33pp; English.
XX
PT   New polypeptides contained in plant virus expression vectors, as gene
PT   expression tools, in protein replacement therapy or for intervening in
PT   a metabolic pathway to improve the nutritional value of a crop or alter
PT   the oil content of seeds -
XX
CC   This invention describes a novel isolated polynucleotide comprising an
CC   internal ribosome entry site (IRES) nucleotide sequence, an open reading
CC   frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,
CC   where the IRES is located between ORF1 and ORF2. The novel
CC   polynucleotides of the invention are used in the construction of a
CC   recombinant potato virus X-based viral vector containing a nucleic acid
CC   construct comprising a bicistronic message with an intervening IRES. The
CC   constructs are used in a method for regulating the rate at which a virus
CC   infection spreads in a host. Regulation is achieved by placing the
CC   nucleic acid construct comprising an internal ribosome entry site
CC   upstream of a coat protein gene, where the IRES is chosen by the rate of
CC   infection of the viral vector on a host in the presence of that IRES. The
CC   polynucleotide and vectors of the invention are useful for directing
CC   rapid and high-level expression of foreign genes in mature,
CC   differentiated, plant tissue. These are particularly useful in protein
CC   replacement therapy, or for intervening in a metabolic pathway to improve
CC   the nutritional value of a crop or alter the oil content of the seed.
CC   This sequence represents a fragment of the clone TXS.GFP.HIRES-CP
CC   containing an IRES region described in the method of the invention.
XX
SQ   Sequence 251 BP; 72 A; 51 C; 76 G; 52 T; 0 other;

Query Match   79.3%; Score 149; DB 24; Length 251;
Best Local Similarity   100.0%;
Matches 149; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 CGAATTCGTCGATTCGGTTGCAGCATTTAAAGCGGTTGACAACTTTAAAGGAAGGAAGAAAGA 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97 CGAATTCGTCGATTCGGTTGCAGCATTTAAAGCGGTTGACAACTTTAAAGGAAGGAAGAAAGA 156

Qy   87 GAAGGTTGAAGAAAAGGGTGTAGTAAGTAAGTATAAGTACAGACCGGAGAAGTACGCCGG 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157 GAAGGTTGAAGAAAAGGGTGTAGTAAGTAAGTATAAGTACAGACCGGAGAAGTACGCCGG 216

Qy   147 TCCTGATTCGTTTAATTTGAAAGAAGAAA 175
              |||||||||||||||||||||||||||||
Db   217 TCCTGATTCGTTTAATTTGAAAGAAGAAA 245

RESULT 5
ID   ABQ76063 standard; DNA; 180 BP.
XX
AC   ABQ76063;
XX
DT   30-SEP-2002 (first entry)
XX
DE   Clone SC196 containing IREScp DNA sequence.
XX
KW   IRES; internal ribosome entry site; viral protein; bicistronic;
KW   virus infection; coat protein; protein replacement therapy; crop;

Left column

```
KW   nutritional value; seed oil content; ds.
XX
OS   Synthetic.
XX
PN   WO20025719-A2.
XX
PF   WO20025719-A2.
XX
PR   09-JAN-2001; 2001US-0758962.
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2002; 2002WO-US01123.
XX
PD   09-JAN-2002; 2002WO-US01123.
XX
PA   (BIOS-) BIOSOURCE GENETICS CORP.
XX
PR   09-JAN-2001; 2001US-0758962.
XX
PI   Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX
DR   WPI; 2002-557829/59.
XX
PT   New polypeptides contained in plant virus expression vectors, as gene
PT   expression tools, in protein replacement therapy or for intervening in
PT   a metabolic pathway to improve the nutritional value of a crop or alter
PT   the oil content of seeds .
XX
PS   Claim 4; Fig 4; 33pp; English.
XX
CC   This invention describes a novel isolated polynucleotide comprising an
CC   internal ribosome entry site (IRES) nucleotide sequence, an open reading
CC   frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,
CC   where the IRES is located between ORF1 and ORF2. The novel
CC   polynucleotides of the invention are used in the construction of a
CC   recombinant potato virus X-based viral vector containing a nucleic acid
CC   construct comprising a bicistronic message with an intervening IRES. The
CC   constructs are used in a method for regulating the rate at which a virus
CC   infection spreads in a host. Regulation is achieved by placing the
CC   nucleic acid construct comprising an internal ribosomes entry site
CC   upstream of a coat protein gene, where the IRES is chosen by the rate of
CC   infection of the viral vector on a host in the presence of that IRES. The
CC   polynucleotide and vectors of the invention are useful for directing
CC   rapid and high-level expression of foreign genes in mature,
CC   differentiated, plant tissue. These are particularly useful in protein
CC   replacement therapy, or for intervening in a metabolic pathway to improve
CC   the nutritional value of a crop or alter the oil content of the seed.
CC   This sequence represents a fragment of the clone SC196 containing an
CC   IREScp region described in the method of the invention.
XX
SQ   Sequence 180 BP; 60 A; 31 C; 52 G; 37 T; 0 other;

Query Match             23.4%;  Score 44;  DB 24;  Length 180;
Best Local Similarity   100.0%;  Pred. No. 9.3e-12;
Matches    44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          90 GGTTGAAGAAAAAGGGTGTAGTAAGTAAGTAACAGACACCGG 133
               |||||||||||||||||||||||||||||||||||||||||||
Db          89 GGTTGAAGAAAAAGGGTGTAGTAAGTAAGTAACAGACACCGG 132


RESULT 6
ABQ76064
ID   ABQ76064 standard; DNA; 181 BP.
XX
AC   ABQ76064;
XX
DT   30-SEP-2002  (first entry)
XX
DE   Clone SC197 containing IREScp DNA sequence.
XX
KW   IRES; internal ribosome entry site; viral protein; bicistronic;
KW   virus infection; coat protein; protein replacement therapy;
KW   nutritional value; seed oil content; ds.
XX
OS   Synthetic.
XX
PN   WO20025719-A2.
```

Right column

```
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2002; 2002WO-US01123.
XX
PR   09-JAN-2001; 2001US-0758962.
XX
PA   (BIOS-) BIOSOURCE GENETICS CORP.
XX
PI   Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX
DR   WPI; 2002-557829/59.
XX
PT   New polypeptides contained in plant virus expression vectors, as gene
PT   expression tools, in protein replacement therapy or for intervening in
PT   a metabolic pathway to improve the nutritional value of a crop or alter
PT   the oil content of seeds .
XX
PS   Claim 4; Fig 4; 33pp; English.
XX
CC   This invention describes a novel isolated polynucleotide comprising an
CC   internal ribosome entry site (IRES) nucleotide sequence, an open reading
CC   frame (ORF1) encoding a peptide and an ORF2 encoding a viral protein,
CC   where the IRES is located between ORF1 and ORF2. The novel
CC   polynucleotides of the invention are used in the construction of a
CC   recombinant potato virus X-based viral vector containing a nucleic acid
CC   construct comprising a bicistronic message with an intervening IRES. The
CC   constructs are used in a method for regulating the rate at which a virus
CC   infection spreads in a host. Regulation is achieved by placing the
CC   nucleic acid construct comprising an internal ribosomes entry site
CC   upstream of a coat protein gene, where the IRES is chosen by the rate of
CC   infection of the viral vector on a host in the presence of that IRES. The
CC   polynucleotide and vectors of the invention are useful for directing
CC   rapid and high-level expression of foreign genes in mature,
CC   differentiated, plant tissue. These are particularly useful in protein
CC   replacement therapy, or for intervening in a metabolic pathway to improve
CC   the nutritional value of a crop or alter the oil content of the seed.
CC   This sequence represents a fragment of the clone SC197 containing an
CC   IREScp region described in the method of the invention.
XX
SQ   Sequence 181 BP; 61 A; 28 C; 53 G; 39 T; 0 other;

Query Match             22.9%;  Score 43;  DB 24;  Length 181;
Best Local Similarity   100.0%;  Pred. No. 2.8e-11;
Matches    43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTCGAATTCGG 43
               ||||||||||||||||||||||||||||||||||||||||||||
Db           1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTCGAATTCGG 43


RESULT 7
ABQ76061
ID   ABQ76061 standard; DNA; 122 BP.
XX
AC   ABQ76061;
XX
DT   30-SEP-2002  (first entry)
XX
DE   Clone TXS.GFP-IRESmp-CP containing IRES DNA sequence.
XX
KW   IRES; internal ribosome entry site; viral protein; bicistronic;
KW   virus infection; coat protein; protein replacement therapy; crop;
KW   nutritional value; seed oil content; ds.
XX
OS   Synthetic.
XX
PN   WO20025719-A2.
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2002; 2002WO-US01123.
```

PR 09-JAN-2001; 2001US-0758962.
XX
PA (BIOS-) BIOSOURCE GENETICS CORP.
XX
PI Santa-Cruz S, Pogue GP, Toth RL, Chapman S, Carr F;
XX
DR WPI; 2002-557829/59.
XX
PT New polypeptides contained in plant virus expression vectors, as gene expression tools, in protein replacement therapy or for intervening in a metabolic pathway to improve the nutritional value of a crop or alter the oil content of seeds
XX
PS Claim 4; Fig 3; 33pp; English.
XX
CC This invention describes a novel isolated polynucleotide comprising an internal ribosome entry site (IRES) nucleotide sequence, an open reading frame (ORF) encoding a peptide and an ORF2 encoding a viral protein, where the IRES is located between ORF1 and ORF2. The novel polynucleotides of the invention are used in the construction of a recombinant potato virus X-based viral vector containing a nucleic acid construct comprising a bicistronic message with an intervening IRES. The constructs are used in a method for regulating the rate at which a virus infection spreads in a host. Regulation is achieved by placing the nucleic acid construct comprising an internal ribosome entry site upstream of a coat protein gene, where the IRES is chosen by the rate of infection of the viral vector on a host in the presence of that IRES. The polynucleotide and vectors of the invention are useful for directing rapid and high-level expression of foreign genes in mature, differentiated, plant tissue. These are particularly useful in a replacement therapy, or for intervening in a metabolic pathway to improve the nutritional value of a crop or alter the oil content of the seed. This sequence represents a fragment of the clone TXS.GFP.IRESmp-CP containing an IRES region described in the method of the invention.
XX
SQ Sequence 122 BP; 30 A; 14 C; 28 G; 50 T; 0 other;

Query Match 18.6%; Score 35; DB 24; Length 122;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
      |||||||||||||||||||||||||||||||||||
Db 1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35

RESULT 8
ID ABV36222/c
XX
AC ABV36222 standard; cDNA; 444 BP.
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 36213.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 7497-7498; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213), of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 444 BP; 91 A; 124 C; 113 G; 116 T; 0 other;

Query Match 18.6%; Score 35; DB 23; Length 444;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
      |||||||||||||||||||||||||||||||||||
Db 268 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 234

RESULT 9
ID ABV45268/c
XX
AC ABV45268 standard; cDNA; 444 BP.
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 45259.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.

```
RESULT 10
ID   ABV49873/c
     ABV49873 standard; cDNA; 514 BP.
XX
AC   ABV49873;
XX
DT   17-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 49864.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US05171.
XX
PR   17-FEB-2000; 2000US-183319P.
PR   16-MAR-2000; 2000US-189862P.
PR   25-MAY-2000; 2000US-207454P.
PR   09-JUN-2000; 2000US-211314P.
PR   18-JUL-2000; 2000US-219007P.
PR   13-DEC-2000; 2000US-255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer
XX
PS   Claim 1; Page 8960; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for:
CC   (a) assessing whether a patient is afflicted with prostate cancer;
CC   (b) monitoring the progression of prostate cancer in a patient;
CC   (c) assessing the efficacy of a test compound to inhibit prostate
CC   cancer in a patient;
CC   (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC   in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound;
CC   (g) determining whether prostate cancer has metastasized in a patient;
CC   (h) assessing the aggressiveness or indolence of prostate cancer in a
CC   patient;
CC   (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ   Sequence 444 BP; 91 A; 124 C; 113 G; 116 T; 0 other;

Query Match     18.6%;  Score 35;  DB 23;  Length 444;
Best Local Similarity   100.0%;   Pred. No. 2e-07;
Matches   35;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
        |||||||||||||||||||||||||||||||||||
Db  268 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 234


RESULT 11
ID   ABK52453
     ABK52453 standard; cDNA; 1353 BP.
XX
AC   ABK52453;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Accessory factor TIP30-32.12 cDNA.
XX
KW   Accessory factor TIP30-32.12; regulator protein; HIV-1 Tat; Immunopathy;
KW   cancer; human immunodeficiency virus; gene; ss.
XX
OS   Unidentified.
XX
FH   Key            Location/Qualifiers
FT   CDS            124..1002
FT                  /*tag= a
FT                  /product= "Accessory factor TIP30-32.12"
XX
PN   CN1331116-A.
XX
PD   16-JAN-2002.
XX
PF   30-JUN-2000; 2000CN-0116960.
XX
PR   30-JUN-2000; 2000CN-0116960.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-316398/36.
DR   P-PSDB; AAU79300.
XX
PT   Accessory factor TIP30-32.12 polypeptide for mutual action of regulator
PT   proteins HIV-1 Tat and polynucleotide encoding it, for treating e.g.
PT   cancer and immunopathy -
XX
PS   Claim 6; Page 27-28 (Disclosure); 35pp; Chinese.
CC
CC   The invention relates to the accessory factor TIP30-32.12 polypeptide for
CC   mutual action of regulator proteins HIV-1 Tat, and its associated
CC   polynucleotide. The polynucleotide, polypeptide and its antagonist are
CC   useful for treating cancer and immunopathy. This sequence represents cDNA
XX
SQ   Sequence 514 BP; 108 A; 123 C; 125 G; 158 T; 0 other;

Query Match     18.6%;  Score 35;  DB 23;  Length 514;
Best Local Similarity   100.0%;   Pred. No. 2e-07;
Matches   35;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
        |||||||||||||||||||||||||||||||||||
Db  347 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 313
```

CC   encoding the accessory factor TIP30-32.12.
XX
SQ   Sequence 1353 BP; 412 A; 329 C; 294 G; 318 T; 0 other;

Query Match   18.6%; Score 35; DB 24; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches   35; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy  1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
       |||||||||||||||||||||||||||||||||||
Db  4 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 38
```

RESULT 12
ID   AAV41999 standard; DNA; 2126 BP.
XX
AC   AAV41999;
XX
DT   20-NOV-1998 (first entry)
XX
DE   Human GPR14 (G-protein coupled receptor polypeptide) gene.
XX
KW   G-protein coupled receptor polypeptide; GPR14; human; fungal infection;
KW   bacterial infection; protozoan infection; viral infection; agonist;
KW   antagonist; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   694..1863
FT   /*tag= a
FT   /product= "human GPR14"
XX
PN   EP859052-A1.
XX
PD   19-AUG-1998.
XX
PF   18-NOV-1997; 97EP-0309252.
PR   27-JAN-1997; 97US-0789354.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI   Bergsma DJ, Shabon U;
XX
DR   WPI; 1998-429680/37.
DR   P-PSDB; AAW59908.
XX
PT   New DNA sequence encoding G-protein coupled receptor polypeptide
PT   GPR14 - and corresponding polypeptide, agonists, antagonists, etc.
XX
PS   Claim 3; Fig 1; 23pp; English.
XX
CC   This is the nucleotide sequence of the gene encoding the novel G-protein
CC   coupled receptor polypeptide (GPR14), used in the method of the
CC   invention. Human GPR14 polypeptides and polynucleotides can be used in
CC   the treatment of infections such as bacterial, fungal, protozoan and
CC   viral infections. Agonists and antagonists can be used to treat
CC   conditions associated with Human GPR14 imbalance.
XX
SQ   Sequence 2126 BP; 317 A; 763 C; 669 G; 377 T; 0 other;

Query Match   18.6%; Score 35; DB 19; Length 2126;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches   35; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
        |||||||||||||||||||||||||||||||||||
Db  30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

RESULT 13

AAZ11174
ID   AAZ11174 standard; cDNA; 2126 BP.
XX
AC   AAZ11174;
XX
DT   04-NOV-1999 (first entry)
XX
DE   Human GPR14 coding sequence.
XX
KW   GPR14; G protein-coupled receptor; ischaemic coronary artery disease;
KW   atherosclerosis; metabolic disease; CHF/myocardial dysfunction; migraine;
KW   arrhythmia; restenosis; hypertension; pulmonary disease;
KW   fibrotic vasculopathy; cerebrovascular event; neurogenic inflammation;
KW   haematopoietic disorder; adult respiratory distress syndrome; ARDS;
KW   cancer; autoimmune disease; therapy; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   694..1863
FT   /*tag= a
FT   /product= GPR14
XX
PN   WO9940192-A1.
XX
PD   12-AUG-1999.
XX
PF   27-JAN-1999; 99WO-US01634.
XX
PR   15-JAN-1999; 99US-0232857.
PR   09-FEB-1998; 98US-0074075.
PR   10-APR-1998; 98US-0058725.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
PA   (SMIK) SMITHKLINE BEECHAM LAB PHARM.
PA   (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI   Aiyar NV, Al-Barazanji K, Ames RS, Arnold AR, Bergsma DJ;
PI   Chambers J, Douglas SA, Foley JJ, Gout B, Khandoudi N;
PI   Sarau HM, Shabon U, Willette RN;
XX
DR   WPI; 1999-527305/44.
DR   P-PSDB; AAY32920.
XX
PT   Human G protein-coupled receptor GPR14, useful for identifying
PT   agonists and antagonists
XX
PS   Claim 4; Fig 1; 64pp; English.
XX
CC   This sequence encodes the human G protein-coupled receptor, GPR14, of the
CC   invention. Human GPR14, polypeptide and polynucleotides are useful in
CC   methods for treatment of, e.g. ischaemic coronary artery disease,
CC   atherosclerosis, metabolic diseases, CHF/myocardial dysfunction,
CC   arrhythmias, restenosis, hypertension and hypotension, pulmonary disease,
CC   fibrotic vasculopathies, cerebrovascular events, neurogenic
CC   inflammation/migraine, haematopoietic disorders, adult respiratory
CC   distress syndrome (ARDS); cancer, autoimmune diseases, etc. The methods
CC   can be used to identify agonists and antagonists of human and rat GPR14.
XX
SQ   Sequence 2126 BP; 317 A; 763 C; 669 G; 377 T; 0 other;

Query Match   18.6%; Score 35; DB 20; Length 2126;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches   35; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
        |||||||||||||||||||||||||||||||||||
Db  30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

RESULT 14
AAK99298
ID   AAK99298 standard; cDNA; 3236 BP.

```
XX
AC   AAK99298;
XX
DT   11-JUN-2002 (first entry)
XX
DE   cDNA of the human cancer suppressor gene 98.
XX
KW   Cytostatic; anti-HIV; human; cancer suppressor gene 98; antagonist;
KW   DNA recombination; cancer; HIV infection; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             76..2760
FT                   /*tag= a
FT                   /product= "Protein of human cancer suppressor gene 98"
XX
PN   CN1328030-A.
XX
PD   26-DEC-2001.
XX
PF   12-JUN-2000; 2000CN-0116437.
XX
PR   12-JUN-2000; 2000CN-0116437.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y, Xie Y;
XX
DR   WPI; 2002-258228/31.
DR   P-PSDB; AAO20441.
XX
PS   Claim 6; Page 25-27 Disclosure; 37pp; Chinese.
XX
PT   Polypeptide-human cancer suppressor gene 98, useful for treating
PT   diseases such as cancer and HIV infection -
XX
CC   The invention relates to a novel protein of the human cancer suppressor
CC   gene 98, the encoding polynucleotide, a process for preparing the
CC   polynucleotide by DNA recombination, and an antagonist to the
CC   polynucleotide. The protein is useful for treating diseases such as
CC   cancer, HIV infection, etc. This polynucleotide sequence represents the
CC   cDNA of the human cancer suppressor gene 98 of the invention.
XX
SQ   Sequence 3236 BP; 892 A; 638 C; 745 G; 961 T; 0 other;

Query Match       18.6%;   Score 35;   DB 24;   Length 3236;
Best Local Similarity    100.0%;   Pred. No. 1.8e-07;
Matches  35;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
         |||||||||||||||||||||||||||||||||||
Db   31 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 65
```

```
XX
PN   CN1331139-A.
XX
PD   16-JAN-2002.
XX
PD   26-JUN-2000; 2000CN-0116749.
XX
PF   26-JUN-2000; 2000CN-0116749.
XX
PR   26-JUN-2000; 2000CN-0116749.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y, Xie Y;
XX
DR   WPI; 2002-292849/34.
DR   P-PSDB; ABB81691.
XX
PS   Claim 6; Page 25-27; 34pp; Chinese.
XX
PT   Polypeptide-human X2 theca repressor 10.34 and polynucleotide for
PT   coding it -
XX
CC   The sequence encodes the novel human X2 theca repressor 10.34 polypeptide
CC   of the invention. The polypeptide is useful for treating diseases such as
CC   immunopathy. The antagonist of the polypeptide and its medical action,
CC   and the application of the polynucleotide are also disclosed.
XX
SQ   Sequence 3905 BP; 1245 A; 731 C; 701 G; 1228 T; 0 other;

Query Match       18.6%;   Score 35;   DB 24;   Length 3905;
Best Local Similarity    100.0%;   Pred. No. 1.8e-07;
Matches  35;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
         |||||||||||||||||||||||||||||||||||
Db   31 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 65
```

Search completed: February 19, 2003, 07:13:20
Job time : 231 secs

```
RESULT 15
ABN83030
ID   ABN83030 standard; cDNA; 3905 BP.
XX
AC   ABN83030;
XX
DT   14-AUG-2002 (first entry)
XX
DE   Human X2 theca repressor 10.34 cDNA.
XX
KW   Human; X2 theca repressor 10.34; repressor; immunopathy; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3574..3858
FT                   /*tag= a
FT                   /product= "X2 theca repressor 10.34"
```

OM nucleic - nucleic search, using sw model

                              GenCore version 5.1.3
                              Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: February 19, 2003, 07:08:12 ; Search time 45 seconds
                                              (without alignments)
                                              1281.227 Million cell updates/sec

Title: US-09-758-962-1
Perfect score: 188
Sequence: 1 gtcgacggtatcgataagct..........gaagaaagagctcaccatgg 188

Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0
Scoring table: Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : , 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   Issued_Patents_NA:*
                1: /cgn2_6/ptodata/2/1na/5A_COMB.seq:*
                2: /cgn2_6/ptodata/2/1na/5B_COMB.seq:*
                3: /cgn2_6/ptodata/2/1na/6A_COMB.seq:*
                4: /cgn2_6/ptodata/2/1na/6B_COMB.seq:*
                5: /cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:*
                6: /cgn2_6/ptodata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | | % | Query | | | | |
No.	Score	Match	Length	DB	ID	Description
1	35	18.6	2126	2	US-08-789-354-1	Sequence 1, Appl1
2	35	18.6	2126	3	US-09-110-937-1	Sequence 1, Appl1
3	35	18.6	2126	3	US-09-058-725B-1	Sequence 1, Appl1
4	35	18.6	2126	4	US-09-233-857-1	Sequence 1, Appl1
5	35	18.6	7560	4	US-09-103-478-4	Sequence 4, Appl1
6	35	18.6	7560	4	US-09-193-931C-4	Sequence 4, Appl1
7	34	18.1	924	5	US-08-983-409-3	Sequence 3, Appl1
8	34	18.1	1270	5	PCT-US95-10403-1	Sequence 1, Appl1
9	34	18.1	1318	5	US-09-183-253-3	Sequence 3, Appl1
10	34	18.1	1471	4	US-09-049-985-11	Sequence 11, Appl1
11	34	18.1	1630	3	US-08-838-219B-8	Sequence 8, Appl1
12	34	18.1	1638	3	US-09-233-336A-8	Sequence 8, Appl1
13	34	18.1	1638	4	US-09-233-752A-8	Sequence 8, Appl1
14	34	18.1	1638	4	US-09-402-036-8	Sequence 8, Appl1
15	34	18.1	1638	4	US-09-904-226-8	Sequence 8, Appl1
16	34	18.1	1679	5	US-08-676-882-1	Sequence 1, Appl1
17	34	18.1	1703	5	PCT-US95-10403-3	Sequence 3, Appl1
18	34	18.1	1958	4	US-09-215-221-9	Sequence 9, Appl1
19	34	18.1	2031	4	US-09-149-476-155	Sequence 155, App
20	34	18.1	2185	3	US-08-467-948A-3	Sequence 3, Appl1
21	34	18.1	2185	4	US-08-467-947A-3	Sequence 3, Appl1
22	34	18.1	2200	5	US-08-680-395-4	Sequence 4, Appl1
23	34	18.1	2200	1	PCT-US95-08565-21	Sequence 21, Appl1
24	34	18.1	2605	2	US-08-272-255-21	Sequence 21, Appl1
25	34	18.1	3344	1	US-07-718-575-13	Sequence 13, Appl1
26	34	18.1	3344	1	US-08-481-206-13	Sequence 13, Appl1
27	34	18.1	3344	2	US-08-486-269A-13	Sequence 13, Appl1
28	34	18.1	5470	2	US-08-443-861-1	Sequence 1, Appl1
29	34	18.1	5470	4	US-08-193-829B-1	Sequence 1, Appl1
30	33	17.6	57	4	US-09-203-681-4	Sequence 4, Appl1
31	33	17.6	78	4	US-09-462-645C-25	Sequence 25, Appl
32	33	17.6	84	4	US-09-462-645C-26	Sequence 26, Appl
33	33	17.6	88	1	US-08-144-602B-15	Sequence 15, Appl1
34	33	17.6	129	3	US-08-752-238-4	Sequence 4, Appl1
35	33	17.6	129	3	US-09-085-603B-4	Sequence 4, Appl1
36	33	17.6	129	4	US-09-085-603B-21	Sequence 21, Appl1
37	33	17.6	129	3	US-09-085-603B-22	Sequence 22, Appl1
38	33	17.6	147	5	US-09-323-872A-26	Sequence 26, Appl1
39	33	17.6	229	4	US-09-116-492A-39	Sequence 39, Appl
40	33	17.6	332	4	US-08-387-707-14	Sequence 14, Appl1
41	33	17.6	332	4	US-08-405-271A-14	Sequence 14, Appl
42	33	17.6	457	4	US-09-643-597-306	Sequence 306, App
43	33	17.6	516	3	US-08-928-799A-4	Sequence 4, Appl1
44	33	17.6	545	4	US-08-991-789A-10	Sequence 10, Appl1
45	33	17.6	545	4	US-09-062-451-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-789-354-1
Sequence 1, Application US/08789354
Patent No. 5851798
GENERAL INFORMATION:
;   APPLICANT: Shabon, Usman
;   APPLICANT: Bergsma, Derk
;   TITLE OF INVENTION: Cloning of Human GPR14 Re
;   TITLE OF INVENTION: ceptor
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Smithkline Beecham Corporation
;   STREET: 709 Swedeland Road
;   CITY: King of Prussia
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19406
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/789,354
;   FILING DATE: 27-JAN-1997
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Han, William T
;   REGISTRATION NUMBER: 34,344
;   REFERENCE/DOCKET NUMBER: P50610
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 610-270-5219
;   TELEFAX: 610-270-4026
;   TELEX:
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2126 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-789-354-1

Query Match   18.6%;   Score 35;   DB 2;   Length 2126;
Best Local Similarity   100.0%;   Pred. No. 2e-08;
Matches   35;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
          |||||||||||||||||||||||||||||||||||
Db     30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

RESULT 2
US-09-110-937-1
Sequence 1, Application US/0911093 7A
Patent No. 6005074
GENERAL INFORMATION:
   APPLICANT: SHABON, USMAN
   APPLICANT: BERGSMA, DERK
   TITLE OF INVENTION: CLONING OF HUMAN GPR14 RECEPTOR
   FILE REFERENCE: P50610-1
   CURRENT APPLICATION NUMBER: US/09/110,937A
   CURRENT FILING DATE: 1998-07-06
   NUMBER OF SEQ ID NOS: 2
   SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 1
   LENGTH: 2126
   TYPE: DNA
   ORGANISM: HOMO SAPIENS
US-09-110-937-1

```
Query Match             18.6%;  Score 35;  DB 3;  Length 2126;
Best Local Similarity  100.0%;  Pred. No. 2e-08;
Matches   35;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
          |||||||||||||||||||||||||||||||||||
Db     30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

RESULT 3
US-09-058-725B-1
Sequence 1, Application US/0905872 5B
Patent No. 6133420
GENERAL INFORMATION:
   APPLICANT: Ames, Robert
   APPLICANT: Sarau, Henry
   APPLICANT: Foley, James
   APPLICANT: Chamber, Jon
   TITLE OF INVENTION: A Method of Finding Angonist
   TITLE OF INVENTION: and Antagonist To Human and Rat GPR 14
   NUMBER OF SEQUENCES: 6
   CORRESPONDENCE ADDRESS:
   ADDRESSEE: Smithkline Beecham Corporation
   STREET: 709 Swedeland Road
   CITY: King of Prussia
   STATE: PA
   COUNTRY: USA
   ZIP: 19406
   COMPUTER READABLE FORM:
   MEDIUM TYPE: Diskette
   COMPUTER: IBM Compatible
   OPERATING SYSTEM: DOS
   SOFTWARE: FastSEQ for Windows Version 2.0
   CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/09/058,725B
   FILING DATE: April 10, 1998
   CLASSIFICATION: 435
   PRIOR APPLICATION DATA:
   APPLICATION NUMBER: 08/789,354
   FILING DATE: 27-JAN-1997
   ATTORNEY/AGENT INFORMATION:
   NAME: Han, William T.
   REGISTRATION NUMBER: 34,344
   REFERENCE/DOCKET NUMBER: GP50005-1
   TELECOMMUNICATION INFORMATION:
   TELEPHONE: 610-270-5219
   TELEFAX: 610-270-5090
   TELEX:
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
   LENGTH: 2126 base pairs
   TYPE: nucleic acid
   STRANDEDNESS: single
   TOPOLOGY: linear
   MOLECULE TYPE: cDNA
US-09-058-725B-1

```
Query Match             18.6%;  Score 35;  DB 3;  Length 2126;
Best Local Similarity  100.0%;  Pred. No. 2e-08;
Matches   35;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
          |||||||||||||||||||||||||||||||||||
Db     30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

RESULT 4
US-09-232-857-1
Sequence 1, Application US/09232857
Patent No. 6159700
GENERAL INFORMATION:
   APPLICANT: DOUGLAS, STEPHEN
   APPLICANT: WILLETTE, ROBERT
   APPLICANT: AIYAR, NAMBI
   APPLICANT: ROMANIC, ANNE
   APPLICANT: KHANDOUDI, NASSIRAH
   APPLICANT: GOUT, BERNARD
   APPLICANT: AL-BARAZANJI, KAMAL
   APPLICANT: AMES, ROBERT S.
   APPLICANT: FOLEY, JAMES J.
   APPLICANT: SARAU, HENRY
   APPLICANT: CHAMBERS, JON K.
   APPLICANT: SHABON, USMAN
   APPLICANT: BERGSMA, DERK
   TITLE OF INVENTION: A METHOD OF FINDING AGONIST
   TITLE OF INVENTION: AND ANTAGONIST TO HUMAN AND RAT GPR14
   NUMBER OF SEQUENCES: 6
   CORRESPONDENCE ADDRESS:
   ADDRESSEE: Ratner & Prestia
   STREET: P.O. Box 980
   CITY: Valley Forge
   STATE: PA
   COUNTRY: USA
   ZIP: 19482
   COMPUTER READABLE FORM:
   MEDIUM TYPE: Diskette
   COMPUTER: IBM Compatible
   OPERATING SYSTEM: DOS
   SOFTWARE: FastSEQ for Windows Version 2.0
   CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/09/232,857
   FILING DATE: 15-JAN-1998
   CLASSIFICATION:
   PRIOR APPLICATION DATA:
   APPLICATION NUMBER: 08/789,354
   FILING DATE: 27-JAN-1997
   APPLICATION NUMBER: 60/074,075
   FILING DATE: 09-FEB-1998
   APPLICATION NUMBER: 09/058,725
   FILING DATE: 10-APR-1998
   ATTORNEY/AGENT INFORMATION:
   NAME: Prestia, Paul F
   REGISTRATION NUMBER: 23,031
   REFERENCE/DOCKET NUMBER: GP-50005-2
   TELECOMMUNICATION INFORMATION:
   TELEPHONE: 610-407-0700
   TELEFAX: 610-407-0700
   TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
   LENGTH: 2126 base pairs
   TYPE: nucleic acid

```
;
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;      MOLECULE TYPE:  cDNA
US-09-232-857-1

     Query Match
     Best Local Similarity   18.6%;  Score 35;  DB 3;  Length 2126;
     Matches  35;  Conservative 0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
         |||||||||||||||||||||||||||||||||||
Db    30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64

RESULT 5
US-09-103-478-4
;  Sequence 4, Application US/09103478
;  Patent No. 6235975
;  GENERAL INFORMATION:
;     APPLICANT:  Harada, John
;     APPLICANT:  Lotan, Tamar
;     APPLICANT:  Ohto, Masa-aki
;     APPLICANT:  Goldberg, Robert B.
;     APPLICANT:  Fischer, Robert L.
;     TITLE OF INVENTION:  LEAFY COTYLEDON1 Genes and Their Uses
;     NUMBER OF SEQUENCES:  29
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Townsend and Townsend and Crew LLP
;        STREET:  Two Embarcadero Center, Eighth Floor
;        CITY:  San Francisco
;        STATE:  California
;        COUNTRY:  USA
;        ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM:  PC-DOS/MS-DOS
;        SOFTWARE:  Patentin Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/09/103,478
;        FILING DATE:  24-JUN-1998
;        CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:  US 09/026,221
;        FILING DATE:  19-FEB-1998
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:  US 08/804,534
;        FILING DATE:  21-FEB-1997
;     ATTORNEY/AGENT INFORMATION:
;        NAME:  Einhorn, Gregory P.
;        REGISTRATION NUMBER:  38,440
;        REFERENCE/DOCKET NUMBER:  023070-077611US
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  (415) 576-0200
;        TELEFAX:  (415) 576-0300
;     INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  7560 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-09-103-478-4

     Query Match
     Best Local Similarity   100.0%;  Score 35;  DB 4;  Length 7560;
     Matches  35;  Conservative 0;  Pred. No. 1.8e-08;  Indels  0;  Gaps  0;

Qy     1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
         |||||||||||||||||||||||||||||||||||
Db    53 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 87
```

```
RESULT 6
US-09-193-931C-4
;  Sequence 4, Application US/09193931C
;  Patent No. 6320102
;  GENERAL INFORMATION:
;     APPLICANT:  Harada, John
;     APPLICANT:  Lotan, Tamar
;     APPLICANT:  Ohto, Masa-aki
;     APPLICANT:  Goldberg, Robert B.
;     APPLICANT:  Fischer, Robert L.
;     APPLICANT:  The Regents of the University of California
;     TITLE OF INVENTION:  LEAFY COTYLEDON1 Genes and Their Uses
;     FILE REFERENCE:  023070-077620
;     CURRENT APPLICATION NUMBER:  US/09/193,931C
;     CURRENT FILING DATE:  1998-11-17
;     PRIOR APPLICATION NUMBER:  US 09/103,478
;     PRIOR FILING DATE:  1998-06-24
;     PRIOR APPLICATION NUMBER:  US 09/026,221
;     PRIOR FILING DATE:  1998-02-19
;     PRIOR APPLICATION NUMBER:  US 08/804,534
;     PRIOR FILING DATE:  1997-02-21
;     NUMBER OF SEQ ID NOS:  29
;     SOFTWARE:  Patentin Ver. 2.0
;  SEQ ID NO 4
;     LENGTH:  7560
;     TYPE:  DNA
;     ORGANISM:  Arabidopsis thaliana
;     FEATURE:
;     OTHER INFORMATION:  7.4 kb genomic wild-type fragment containing LEC1
;     OTHER INFORMATION:  gene
;        NAME/KEY:  misc_feature
;        LOCATION:  (2430)..(5824)
;     OTHER INFORMATION:  corresponds to SEQ ID NO:3
;        NAME/KEY:  promoter
;        LOCATION:  (2430)..(4427)
;     OTHER INFORMATION:  corresponds to LEC1 promoter in SEQ ID NO:3
;        NAME/KEY:  gene
;        LOCATION:  (2430)..(5824)
;        NAME/KEY:  CDS
;        LOCATION:  (4427)..(5054)
;     OTHER INFORMATION:  corresponds to LEC1
;        NAME/KEY:  LEAFY COTYLEDON1 (LEC1)
;        LOCATION:  modified_base
;        NAME/KEY:  (1)..(7560)
;     OTHER INFORMATION:  n = g, a, c or t
US-09-193-931C-4

     Query Match
     Best Local Similarity   18.6%;  Score 35;  DB 4;  Length 7560;
     Matches  35;  Conservative 0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
         |||||||||||||||||||||||||||||||||||
Db    53 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 87

RESULT 7
US-08-983-409-3
;  Sequence 3, Application US/08983409
;  Patent No. 6150586
;  GENERAL INFORMATION:
;     APPLICANT:  Slabas, Antoni
;     APPLICANT:  Elborough, Kieran
;     TITLE OF INVENTION:  Plant Gene Encoding Acetyl Coenzyme A
;     TITLE OF INVENTION:  Carboxylase Biotin Carboxyl Carrier Protein
;     NUMBER OF SEQUENCES:  6
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Zeneca Inc.
;        STREET:  1800 Concord Pike
;        CITY:  Wilmington
;        STATE:  DE
;        COUNTRY:  USA
;        ZIP:  19850
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/983,409
FILING DATE: 20-JAN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 06-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/01894
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9516961.1
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hohenschutz, Liza D.
REGISTRATION NUMBER: 33,712
REFERENCE/DOCKET NUMBER: SEE45001/UST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 886-1699
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 924 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-983-409-3

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
        ||||||||||||||||||||||||||||||||||
Db   12 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 45

Query Match   18.1%;   Score 34;   DB 3;   Length 924;
Best Local Similarity   100.0%;   Pred. No. 6.6e-08;
Matches   34;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 8
Sequence 1, Application PCT/US9510403
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: GA4 DNA, Protein and Methods of Use
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein and Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10403
FILING DATE: 15 August 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/291,939
FILING DATE: 16 August 1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0609.408PC00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1270 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: 107..1183
PCT-US95-10403-1

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
        ||||||||||||||||||||||||||||||||||
Db   6 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 39

Query Match   18.1%;   Score 34;   DB 5;   Length 1270;
Best Local Similarity   100.0%;   Pred. No. 6.4e-08;
Matches   34;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 9
US-09-183-253-3/c
Sequence 3, Application US/09183253
Patent No. 6043054
GENERAL INFORMATION:
APPLICANT: VAWTER, LISA
APPLICANT: STAMMERS, MELANIE
TITLE OF INVENTION: NOVEL COMPOUNDS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183,253
FILING DATE: 30-OCT-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9817907.0
FILING DATE: 17-AUG-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/075,306
FILING DATE: 20-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-70395
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1318 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-183-253-3

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
        ||||||||||||||||||||||||||||||||||
Db 1318 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 1285

Query Match   18.1%;   Score 34;   DB 3;   Length 1318;
Best Local Similarity   100.0%;   Pred. No. 6.4e-08;
Matches   34;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 10
US-09-492-985-11/c
Sequence 11, Application US/09492985
Patent No. 6376240
GENERAL INFORMATION:
; APPLICANT: Song, An M.
; APPLICANT: Chen, Ya-Fen
; APPLICANT: Krensky, Alan M.
; TITLE OF INVENTION: RFLAT-1: A Transcription Factor That
; TITLE OF INVENTION: Activates RANTES Gene Expression
; FILE REFERENCE: SUN-113P
; CURRENT APPLICATION NUMBER: US/09/492,985
; CURRENT FILING DATE: 2000-01-27
; EARLIER APPLICATION NUMBER: 60/117,576
; EARLIER FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0.
; SEQ ID NO 11
; LENGTH: 1471
; TYPE: DNA
; ORGANISM: mouse
US-09-492-985-11

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
           ||||||||||||||||||||||||||||||||||
Db   1443 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 1410

Query Match   18.1%; Score 34; DB 4; Length 1471;
Best Local Similarity   100.0%; Pred. No. 6.4e-08;
Matches   34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
US-08-838-219B-8/c
Sequence 8, Application US/08838219B
Patent No. 5877012
GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: A No. 5877012e1 Class of Proteins for the
; TITLE OF INVENTION: Control of Plant Pests
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,219B
; PRIOR APPLICATION DATA:
; CLASSIFICATION: 530
; FILING DATE:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018

; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1925
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1638 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1191
; OTHER INFORMATION: /product= "Translation of cDNA
; OTHER INFORMATION: encoding VIP3A(a) receptor"
US-08-838-219B-8

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
           ||||||||||||||||||||||||||||||||||
Db   1633 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 1600

Query Match   18.1%; Score 34; DB 2; Length 1638;
Best Local Similarity   100.0%; Pred. No. 6.3e-08;
Matches   34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
US-09-233-336A-8/c
Sequence 8, Application US/09233336A
Patent No. 6107279
GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: A No. 6107279e1 Class of Proteins for the
; TITLE OF INVENTION: Control of Plant Pests
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/233,336A
; PRIOR APPLICATION DATA:
; CLASSIFICATION:
; FILING DATE:
; APPLICATION NUMBER: US/09/233,336A
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,219
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/838,219
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594

RESULT 13
US-09-233-752A-8/c
; Sequence 8, Application US/09233752A
; Patent No. 6137033
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: A No. 6137033e1 Class of Proteins for the
; TITLE OF INVENTION: Control of Plant Pests
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; ZIP: 10532
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/233,752A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/838,219

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1925
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8689
; TELEFAX: 919-541-8582
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1638 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1191
; OTHER INFORMATION:
US-09-233-336A-8 OTHER INFORMATION: /product= "Translation of cDNA
;   encoding VIP3A(a) receptor"

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATCG 34
           ||||||||||||||||||||||||||||||||||
Db   1633 GTCGACGGTATCGATAAGCTTGATATCGAATCG 1600

Query Match   18.1%; Score 34; DB 3; Length 1638;
Best Local Similarity 100.0%; Pred. No. 6.3e-08;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1925
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8689
; TELEFAX: 919-541-8582
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1638 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1191
; OTHER INFORMATION:
US-09-233-752A-8 OTHER INFORMATION: /product= "Translation of cDNA
;   encoding VIP3A(a) receptor"

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATCG 34
           ||||||||||||||||||||||||||||||||||
Db   1633 GTCGACGGTATCGATAAGCTTGATATCGAATCG 1600

Query Match   18.1%; Score 34; DB 3; Length 1638;
Best Local Similarity 100.0%; Pred. No. 6.3e-08;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
US-09-402-036-8/c
; Sequence 8, Application US/09402036
; Patent No. 6291156
; GENERAL INFORMATION:
; APPLICANT: Estruch, Juan J.
; APPLICANT: Yu, Cao-Guo
; APPLICANT: Warren, Gregory W.
; APPLICANT: Desai, Nalini
; APPLICANT: Koziel, Michael
; APPLICANT: Nye, Gordon
; TITLE OF INVENTION: Plant Pest Control
; FILE REFERENCE: S-212284C
; CURRENT APPLICATION NUMBER: US/09/402,036
; CURRENT FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: PCT/EP98/01952
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 08/838,219
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/832,263
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/832,265
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/463,483
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 08/314,594
; PRIOR FILING DATE: 1994-09-09
; PRIOR APPLICATION NUMBER: 08/218,018
; PRIOR FILING DATE: 1994-03-24
; PRIOR APPLICATION NUMBER: 08/037,057
; PRIOR FILING DATE: 1993-03-25
; NUMBER OF SEQ ID NOS: 22

```
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 8
;    LENGTH: 1638
;    TYPE: DNA
;    ORGANISM: Agrotis Ipsilon
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (2)..(1189)
;    OTHER INFORMATION:
;    OTHER INFORMATION: Translation of cDNA encoding VIP3A(a) receptor
US-09-402-036-8
```

```
Qy      Query Match          18.1%;   Score 34;   DB 4;   Length 1638;
Db      Best Local Similarity 100.0%;   Pred. No. 6.3e-08;
        Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

        1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
          ||||||||||||||||||||||||||||||||||
     1633 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 1600
```

```
RESULT 15
US-09-904-226-8/c
;  Sequence 8, Application US/09904226
Patent No. 6429360
;  GENERAL INFORMATION:
;    APPLICANT: Estruch, Juan J.
;    APPLICANT: Warren, Gregory W.
;    APPLICANT: Desai, Nalini
;    APPLICANT: Koziel, Michael
;    APPLICANT: Nye, Gordon
;    TITLE OF INVENTION: Plant Pest Control
;    FILE REFERENCE: S-21284D
;    CURRENT APPLICATION NUMBER: US/09/904,226
;    CURRENT FILING DATE: 2001-07-12
;    PRIOR APPLICATION NUMBER: PCT/EP98/01952
;    PRIOR FILING DATE: 1998-04-02
;    PRIOR APPLICATION NUMBER: 08/838,219
;    PRIOR FILING DATE: 1997-04-03
;    PRIOR APPLICATION NUMBER: 08/832,263
;    PRIOR FILING DATE: 1997-04-03
;    PRIOR APPLICATION NUMBER: 08/832,265
;    PRIOR FILING DATE: 1997-04-03
;    PRIOR APPLICATION NUMBER: 08/463,483
;    PRIOR FILING DATE: 1995-06-05
;    PRIOR APPLICATION NUMBER: 08/314,594
;    PRIOR FILING DATE: 1994-09-09
;    PRIOR APPLICATION NUMBER: 08/218,018
;    PRIOR FILING DATE: 1994-03-24
;    PRIOR APPLICATION NUMBER: 08/037,057
;    PRIOR FILING DATE: 1993-03-25
;    NUMBER OF SEQ ID NOS: 22
;    SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 8
;    LENGTH: 1638
;    TYPE: DNA
;    ORGANISM: Agrotis Ipsilon
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (2)..(1189)
;    OTHER INFORMATION: Translation of cDNA encoding VIP3A(a) receptor
US-09-904-226-8
```

```
Qy      Query Match          18.1%;   Score 34;   DB 4;   Length 1638;
Db      Best Local Similarity 100.0%;   Pred. No. 6.3e-08;
        Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

        1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
          ||||||||||||||||||||||||||||||||||
     1633 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 1600
```

```
                                            GenCore version 5.1.3
                                            Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:                February 19, 2003, 07:00:11 ; Search time 1949 seconds
                                                     (without alignments)
                                                     1562.212 Million cell updates/sec

Title:                 US-09-758-962-1
Perfect score:         188
Sequence:              1 gtcgacggtatcgataagct.........gaagaaagagctcaccatgg 188

Scoring table:         OLIGO.NUC
                       Gapop 60.0 , Gapext 60.0

Searched:              16154066 seqs, 8097743376 residues

Word size :            0

Total number of hits satisfying chosen parameters:    32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :     EST:*
          1: em_estba:*
          2: em_esthum:*
          3: em_estin:*
          4: em_estmu:*
          5: em_estov:*
          6: em_estpl:*
          7: em_estro:*
          8: em_htc:*
          9: gb_est1:*
         10: gb_est2:*
         11: gb_htc:*
         12: gb_est3:*
         13: gb_est4:*
         14: gb_est5:*
         15: em_estfun:*
         16: em_estom:*
         17: gb_gss:*
         18: em_gss_hum:*
         19: em_gss_inv:*
         20: em_gss_pln:*
         21: em_gss_vrt:*
         22: em_gss_fun:*
         23: em_gss_mam:*
         24: em_gss_mus:*
         25: em_gss_other:*
         26: em_gss_pro:*
         27: em_gss_rod:*
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	38	20.2	304	9	AU167182	AU167182 AU167182
c 2	36	19.1	818	13	BM141055	BM141055 2FO441 2e
c 3	35	18.6	495	12	BG139700	BG139700 EST480142
4	35	18.6	919	13	BM351825	BM351825 Tgz244h01
5	35	18.6	945	13	BM351824	BM351824 Tgz240h11
6	35	18.6	1300	13	BM351823	BM351823 Tgz240d09

ALIGNMENTS

```
RESULT 1
AU167182
AU167182/c
LOCUS       AU167182                 304 bp    mRNA     linear       EST 29-JAN-2001
DEFINITION  AU167182 Ol-br-ad cDNA Oryzias latipes cDNA clone br0230, mRNA
            sequence.
ACCESSION   AU167182
VERSION     AU167182.1  GI:12589251
KEYWORDS    EST.
SOURCE      Japanese medaka.
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1  (bases 1 to 304)
  AUTHORS   Mita,K., Ishikawa,Y. and Yamauchi,M.
  TITLE     Establishment of cDNA database of medaka, Oryzias latipes
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Mita K
            Genome Research Group
            National Institute of Radiological Sciences
            Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
            Email: kmita@nirs.go.jp
            method:uni-directional sequence direction:sequenced from T3 primer
            (5' -> 3').
FEATURES             Location/Qualifiers
     source          1..304
                     /organism="Oryzias latipes"
                     /organism="Oryzias latipes"
                     /strain="HNI"
```

```
  c  7  34 18.1  52 10 AW172124  AW172124 SWL4CAK07
  c  8  34 18.1  53 10 AW042087  AW042087 SWL4CAK03
  c  9  34 18.1  53 10 AW626527  AW626527 SWOVL3CAN
  c 10  34 18.1  54 10 AW172172  AW172172 SWL4CAK07
  c 11  34 18.1  54 12 BE758345  BE758345 SWYACAL08
  c 12  34 18.1  66 10 AW626545  AW626545 SWOVL3CAN
  c 13  34 18.1  68 10 AW159927  AW159927 SWL4CAK04
  c 14  34 18.1  91  9 AI068878  AI068878 SWBML3SDI
  c 15  34 18.1  92  9 AA680557  AA680557 SWOVL3CAN
  c 16  34 18.1  92 14 BQ821730  BQ821730 1030079G0
  c 17  34 18.1 105 14 BQ819685  BQ819685 1030094E0
  c 18  34 18.1 143  9 AL837413  AL837413 AL837413
  c 19  34 18.1 175  9 AL835751  AL835751 AL835751
  c 20  34 18.1 181 12 BG360460  BG360460 BRI10632
  c 21  34 18.1 186 10 AW740238  AW740238 BRI10260
  c 22  34 18.1 190 10 AW739962  AW739962 BRI00475
  c 23  34 18.1 206 10 AV743048  AV743048 AV743048
  c 24  34 18.1 212 10 AW739871  AW739871 BRI00384
  c 25  34 18.1 215 10 AW739826  AW739826 BRI00339
  c 26  34 18.1 217 10 AW740190  AW740190 BRI10211
  c 27  34 18.1 217 10 AW740458  AW740458 BRI10415
  c 28  34 18.1 217 10 BE530967  BE530967 OC03G04 O
  c 29  34 18.1 218 10 AJ407770  AJ407770 AJ407770
  c 30  34 18.1 223 10 AW739964  AW739964 BRI00477
  c 31  34 18.1 227 10 AV742899  AV742899 AV742899
  c 32  34 18.1 231 10 AW740521  AW740521 BRI10478
  c 33  34 18.1 232 10 AW739775  AW739775 BRI00288
  c 34  34 18.1 233 10 AW740417  AW740417 BRI10560
  c 35  34 18.1 240 10 AW740289  AW740289 BRI10311
  c 36  34 18.1 241 10 AJ477003  AJ477003 AJ477003
  c 37  34 18.1 253 10 AW739737  AW739737 BRI00250
  c 38  34 18.1 254 10 AW740254  AW740254 BRI10276
  c 39  34 18.1 258  9 AW740521  AW740521 BRI10478
  c 40  34 18.1 259 10 AI052896  AI052896 SWBML3SBH
  c 41  34 18.1 259 10 AW739879  AW739879 BRI00392
  c 42  34 18.1 259 10 AW740178  AW740178 BRI10199
  c 43  34 18.1 263 10 AW740378  AW740378 BRI10521
  c 44  34 18.1 263 14 BQ815505  BQ815505 1030050R0
  c 44  34 18.1 267 10 AW739838  AW739838 BRI00351
  c 45  34 18.1 274 13 BI119516  BI119516 ARI8E10 P
```

BASE COUNT 106 a 95 c 54 g 49 t
ORIGIN

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTCGA 38
        ||||||||||||||||||||||||||||||||||||||
Db 304 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTCGA 267

Query Match   20.2%;   Score 38;   DB 9;   Length 304;
Best Local Similarity 100.0%;   Pred. No. 2.9e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
                         /db_xref="taxon:8090"
                         /clone="br0230"
                         /clone_lib="ol-br-ad cDNA"
                         /sex="female/male mixed"
                         /tissue_type="brain"
                         /dev_stage="adult"
```

RESULT 2
LOCUS   BM141055   818 bp   mRNA   linear   EST 30-JUN-2002
DEFINITION   ZFO441 Zebrafish Ovary Danio rerio cDNA 5', mRNA sequence.
ACCESSION   BM141055
VERSION   BM141055.1 GI:21630868
KEYWORDS   EST.
SOURCE   zebrafish.
  ORGANISM   Danio rerio
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
             ; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 818)
  AUTHORS   Zeng,S. and Gong,Z.
  TITLE   Zebrafish EST clones from testis and ovary cDNA libraries
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Zhiyuan Gong
             School of Biological Sciences
             National University of Singapore
             Lower Kent Ridge Road, Singapore 119260
             Tel: 65-8742860
             Fax: 65-7792486
             Email: dbsgzy@leonis.nus.edu.sg
             Seq primer: SK (5'-CGCTCTAGAACTAGGATC).

FEATURES   Location/Qualifiers
     source   1..818
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone_lib="zebrafish Ovary"
                     /sex="female"
                     /dev_stage="4-5 month-old female adult fish"
                     /note="Organ: Ovary; Vector: Lambda Uni-ZAP XR; Site_1:
                     EcoRI; Site_2: XhoI; Poly A+ RNA was isolated from the
                     ovaries of 2 female adult zebrafish of 4-5 month-old and
                     the cDNAs were made using oligo dT primers and inserted
                     into lambda ZAP vector at EcoRI (5'-cDNA) and XhoI
                     (3'-cDNA) sites using Stratagene's Lambda Uni-ZAP XR
                     cloning system."

BASE COUNT 255 a 196 c 175 g 189 t 3 others
ORIGIN

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTC 36
        ||||||||||||||||||||||||||||||||||||
Db 241 GTCGACGGTATCGATAAGCTTGATATCGAATTCGTC 276

Query Match   19.1%;   Score 36;   DB 13;   Length 818;
Best Local Similarity 100.0%;   Pred. No. 2.8e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
LOCUS   BG139700
DEFINITION   EST480142 wild tomato pollen Lycopersicon pennellii cDNA clone
             cLPP13P12 5' sequence, mRNA sequence.

ACCESSION   BG139700
VERSION   BG139700.1 GI:12639888
KEYWORDS   EST.
SOURCE   Lycopersicon pennellii.
  ORGANISM   Lycopersicon pennellii
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
             Lycopersicon.
REFERENCE   1 (bases 1 to 495)
  AUTHORS   van der Hoeven,R., Bezerides,J., Sun,H., Cho,J., Utterback,T.,
             Hansen,C., Ronning,C. and Tanksley,S.
  TITLE   Generation of ESTs from wild tomato (L. pennellii) pollen
  JOURNAL   Unpublished (2001)
COMMENT   Contact: CUGI
             Clemson University Genomics Institute
             Clemson University
             100 Jordan Hall, Clemson, SC 29634, USA
             Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES   Location/Qualifiers
     source   1..495
                     /organism="Lycopersicon pennellii"
                     /cultivar="TA56"
                     /db_xref="taxon:28526"
                     /clone="cLPP13P12"
                     /clone_lib="wild tomato pollen"
                     /tissue_type="pollen"
                     /dev_stage="pollen collected from open flowers"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Pollen was collected from open flowers from
                     L. pennellii TA56, and stored at -80 C until library
                     construction."

BASE COUNT 150 a 110 c 104 g 131 t
ORIGIN

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
        |||||||||||||||||||||||||||||||||||
Db   43 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 9

Query Match   18.6%;   Score 35;   DB 12;   Length 495;
Best Local Similarity 100.0%;   Pred. No. 8.7e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
LOCUS   BM351825   919 bp   mRNA   linear   EST 29-JUL-2002
DEFINITION   Tgzz44h01.s1 TgME49 invivo Bradyzoite cDNA clone Tgzz44h01.s1 5', mRNA sequence.
ACCESSION   BM351825
VERSION   BM351825.1 GI:18084183
KEYWORDS   EST.
SOURCE   Toxoplasma gondii.
  ORGANISM   Toxoplasma gondii
             Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida;
             Sarcocystidae; Toxoplasma.
REFERENCE   1 (bases 1 to 919)
  AUTHORS   Cleary,M.D., Singh,U., Blader,I.J., Brewer,J.L. and Boothroyd,J.C.
  TITLE   Toxoplasma gondii asexual development: identification of
             developmentally regulated genes and distinct patterns of gene
             expression
  JOURNAL   Eukaryotic Cell 1 (3), 329-340 (2002)
COMMENT   Contact: Upinder Singh
             Microbiology and Immunology
             Stanford University
             D-305 Fairchild Building, 300 Pasteur Drive Stanford, CA 94305, USA
             Tel: 650 723 7296
             Fax: 650 723 6853
             Email: usingh@stanford.edu
             Seq primer: T7
             High quality sequence start: 20
             High quality sequence stop: 700

RESULT 5

```
BM351824
LOCUS       BM351824       945 bp    mRNA    linear   EST 29-JUL-2002
DEFINITION  Tgzz40h11.s1 TgME49 invivo Bradyzoite cDNA size selected Toxoplasma
            gondii cDNA clone Tgzz40h11.s1 5', mRNA sequence.
ACCESSION   BM351824
VERSION     BM351824.1   GI:18084182
KEYWORDS    EST.
SOURCE      Toxoplasma gondii.
  ORGANISM  Toxoplasma gondii
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Toxoplasma.
REFERENCE   1 (bases 1 to 945)
  AUTHORS   Cleary,M.D., Singh,U., Blader,I.J., Brewer,J.L. and Boothroyd,J.C.
  TITLE     Toxoplasma gondii asexual development: Identification of
            developmentally regulated genes and distinct patterns of gene
            expression
  JOURNAL   Eukaryotic Cell 1 (3), 329-340 (2002)
COMMENT     Contact: Upinder Singh
            Microbiology and Immunology
            Stanford University
            D-305 Fairchild Building, 300 Pasteur Drive Stanford, CA 94305, USA
            Tel: 650 723 7296
            Fax: 650 723 6853
            Email: usingh@stanford.edu
            Seq primer: T7
            High quality sequence start: 20
            High quality sequence stop: 700
            POLYA=NO.
FEATURES             Location/Qualifiers
     source          1..945
                     /organism="Toxoplasma gondii"
                     /strain="ME49"
                     /db_xref="taxon:5811"
                     /clone="Tgzz40h11.s1"
                     /clone_lib="TgME49 invivo Bradyzoite cDNA size selected"
                     /dev_stage="Bradyzoite"
                     /lab_host="DH10"
                     /note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
```

```
Query Match      18.6%;    Score 35;    DB 13;    Length 919;
Best Local Similarity 100.0%;    Pred. No. 8.6e-07;
Matches  35;    Conservative  0;    Mismatches  0;    Indels  0;    Gaps  0;

Qy    1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
         |||||||||||||||||||||||||||||||||||
Db   30 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 64
```

```
BASE COUNT    217 a    219 c    232 g    231 t    20 others
ORIGIN
```

```
FEATURES             Location/Qualifiers
     source          1..919
                     /organism="Toxoplasma gondii"
                     /strain="ME49"
                     /db_xref="taxon:5811"
                     /clone="Tgzz44h01.s1"
                     /clone_lib="TgME49 invivo Bradyzoite cDNA size selected"
                     /dev_stage="Bradyzoite"
                     /lab_host="DH10"
                     /note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
POLYA=Yes.
```

NotI: Mature bradyzoites were obtained from infected mouse brains by percoll density centrifugation. The original library was constructed by Steve Parmley, Palo Alto Medical Foundation. cDNAs were synthesized by priming with oligo d(T) and directionally cloned into the EcoRI/NotI sites of lambda gt11. Warning: the library contains a small percentage of host cDNAs derived from mouse cells. Inserts from this cDNA library were excised with NotI and EcoRI, size selected in a range of 0.7 - 2.0 kb and subcloned into Bluescript II SK- (Adrian Hehl, Ian Manger and John Boothroyd, Stanford University)"

RESULT 6

```
BM351823
LOCUS       BM351823      1300 bp    mRNA    linear   EST 29-JUL-2002
DEFINITION  Tgzz40d09.s1 TgME49 invivo Bradyzoite cDNA size selected Toxoplasma
            gondii cDNA clone Tgzz40d09.s1 5' end sequenced, mRNA sequence.
ACCESSION   BM351823
VERSION     BM351823.1   GI:18084181
KEYWORDS    EST.
SOURCE      Toxoplasma gondii.
  ORGANISM  Toxoplasma gondii
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Toxoplasma.
REFERENCE   1 (bases 1 to 1300)
  AUTHORS   Cleary,M.D., Singh,U., Blader,I.J., Brewer,J.L. and Boothroyd,J.C.
  TITLE     Toxoplasma gondii asexual development: Identification of
            developmentally regulated genes and distinct patterns of gene
            expression
  JOURNAL   Eukaryotic Cell 1 (3), 329-340 (2002)
COMMENT     Contact: Upinder Singh
            Microbiology and Immunology
            Stanford University
            D-305 Fairchild Building, 300 Pasteur Drive Stanford, CA 94305, USA
            Tel: 650 723 7296
            Fax: 650 723 6853
            Email: usingh@stanford.edu
            Seq primer: T7
            High quality sequence start: 50
            High quality sequence stop: 750
            POLYA=NO.
FEATURES             Location/Qualifiers
     source          1..1300
                     /organism="Toxoplasma gondii"
                     /strain="ME49"
                     /db_xref="taxon:5811"
                     /clone="Tgzz40d09.s1"
                     /clone_lib="TgME49 invivo Bradyzoite cDNA size selected"
                     /dev_stage="Bradyzoite"
                     /lab_host="DH10"
                     /note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
```

```
Query Match      18.6%;    Score 35;    DB 13;    Length 945;
Best Local Similarity 100.0%;    Pred. No. 8.6e-07;
Matches  35;    Conservative  0;    Mismatches  0;    Indels  0;    Gaps  0;

Qy    1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
         |||||||||||||||||||||||||||||||||||
Db   46 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 80
```

```
BASE COUNT    199 a    182 c    266 g    261 t    37 others
ORIGIN
```

```
FEATURES             Location/Qualifiers
     source          1..1300
                     /organism="Toxoplasma gondii"
                     /strain="ME49"
                     /db_xref="taxon:5811"
                     /clone="Tgzz40d09.s1"
                     /clone_lib="TgME49 invivo Bradyzoite cDNA size selected"
                     /dev_stage="Bradyzoite"
                     /lab_host="DH10"
                     /note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
POLYA=NO.
```

NotI: Mature bradyzoites were obtained from infected mouse brains by percoll density centrifugation. The original library was constructed by Steve Parmley, Palo Alto Medical Foundation. cDNAs were synthesized by priming with oligo d(T) and directionally cloned into the EcoRI/NotI sites of lambda gt11. Warning: the library contains a small percentage of host cDNAs derived from mouse cells. Inserts from this cDNA library were excised with NotI and EcoRI, size selected in a range of 0.7 - 2.0 kb and subcloned into Bluescript II SK- (Adrian Hehl, Ian Manger and John Boothroyd, Stanford University)"

RESULT 7
AW172124/c

```
Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 35
       |||||||||||||||||||||||||||||||||||
Db  56 GTCGACGGTATCGATAAGCTTGATATCGAATTCGT 90
```

Query Match
Best Local Similarity　18.6%;　Score 35;　DB 13;　Length 1300;
　　　　　　　　　　　100.0%;　Pred. No. 8.6e-07;
Matches　35;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

BASE COUNT　　　284 a　296 c　316 g　290 t　114 others
ORIGIN

```
LOCUS       AW172124                52 bp    mRNA    linear   EST 15-NOV-1999
DEFINITION  SWL4CAK07B01SK Brugia malayi L4 cDNA (SAW99MLW-BmL4) Brugia malayi
            cDNA clone SWL4CAK07B01 5', mRNA sequence.
ACCESSION   AW172124
VERSION     AW172124.1  GI:6431920
KEYWORDS    EST.
SOURCE      Brugia malayi.
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in fourth stage larvae of Brugia malayi
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1..52
                     /organism="Brugia malayi"
                     /db_xref="taxon:6279"
                     /clone="SWL4CAK07B01"
                     /clone_lib="Brugia malayi L4 cDNA (SAW99MLW-BmL4)"
                     /dev_stage="larval stage four"
                     /lab_host="Xl1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Lymphatic filarial nematode parasite of humans.
                     mRNA was prepared from L4s isolated from the peritoneal
                     cavity of jirds and converted to double-stranded cDNA
                     using reverse transcriptase and oligo(dT) followed by
                     RNase H and DNA pol I. The library was constructed by
                     independent recombinants and the average insert size is
                     approx. 1050bp. The library is available from Dr. S.A.
                     Lizotte-Waniewski. The library is available from Dr. S.A.
                     Williams, email:genome@neal.smith.edu."
```

RESULT 8
AW042087/c

```
Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
       ||||||||||||||||||||||||||||||||||
Db  43 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 10
```

Query Match
Best Local Similarity　18.1%;　Score 34;　DB 10;　Length 52;
　　　　　　　　　　　100.0%;　Pred. No. 2.8e-06;
Matches　34;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

BASE COUNT　　　11 a　15 c　11 g　15 t
ORIGIN

```
LOCUS       AW042087                53 bp    mRNA    linear   EST 15-SEP-1999
DEFINITION  SWL4CAK03G05SK Brugia malayi L4 cDNA (SAW99MLW-BmL4) Brugia malayi
            cDNA clone SWL4CAK03G05 5', mRNA sequence.
ACCESSION   AW042087
VERSION     AW042087.1  GI:5900919
KEYWORDS    EST.
SOURCE      Brugia malayi.
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in fourth stage larvae of Brugia malayi
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1..53
                     /organism="Brugia malayi"
                     /db_xref="taxon:6279"
                     /clone="SWL4CAK03G05"
                     /clone_lib="Brugia malayi L4 cDNA (SAW99MLW-BmL4)"
                     /dev_stage="larval stage four"
                     /lab_host="Xl1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Lymphatic filarial nematode parasite of humans.
                     mRNA was prepared from L4s isolated from the peritoneal
                     cavity of jirds and converted to double-stranded cDNA
                     using reverse transcriptase and oligo(dT) followed by
                     RNase H and DNA pol I. The library was constructed by
                     independent recombinants and the average insert size is
                     approx. 1050bp. The library is available from Dr. S.A.
                     Lizotte-Waniewski. The library is available from Dr. S.A.
                     Williams, email:genome@neal.smith.edu."
```

RESULT 9
AW626527/c

```
Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
       ||||||||||||||||||||||||||||||||||
Db  43 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 10
```

Query Match
Best Local Similarity　18.1%;　Score 34;　DB 10;　Length 53;
　　　　　　　　　　　100.0%;　Pred. No. 2.8e-06;
Matches　34;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

BASE COUNT　　　11 a　15 c　11 g　16 t
ORIGIN

```
LOCUS       AW626527                53 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  SWOVL3CAN65A07SK Onchocerca volvulus infective larva cDNA
            (SAW94WL-OvL3) Onchocerca volvulus cDNA clone SWOVL3CAN65A07 5',
            mRNA sequence.
ACCESSION   AW626527
VERSION     AW626527.1  GI:7342392
KEYWORDS    EST.
SOURCE      Onchocerca volvulus.
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Williams,S.A., Lu,W., Lizotte-Waniewski,M. and Laney,S.J.
  TITLE     Genes expressed in infective third stage larvae of Onchocerca
            volvulus
  JOURNAL   Unpublished (1995)
  COMMENT   Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
```

Column 1

```
                    Fax: 413853786
                    Email: genome@smith.edu
                    Seq primer: pBluescript SK.
FEATURES
    source          Location/Qualifiers
                    1. .53
                    /organism="Onchocerca volvulus"
                    /strain="Sierra Leone"
                    /db_xref="taxon:6282"
                    /clone="SWOv3CAN65A07"
                    /clone_lib="Onchocerca volvulus infective larva cDNA
                    (SAW94WL-OvL3)"
                    /lab_host="XL1-Blue MRF'"
                    /note="Vector: lambda Unizap XR; Site_1: EcoR I; Site_2:
                    Xho I; Cutaneous filarial nematode parasite of humans.
                    mRNA was prepared from third stage infective larvae of
                    Onchocerca volvulus isolated from mosquitoes 10 days after
                    infection and converted to double stranded cDNA using
                    reverse transcriptase and oligo(dT) followed by RNase H
                    and DNApol I. The library had 1.8 x 105 independent
                    recombinants and average insert size was 900 base pairs.
                    The library was constructed by Wenhong Lu. The library is
                    available from Dr. S.A. Williams, email genome@smith.edu."
BASE COUNT         11 a     15 c     11 g     16 t
ORIGIN

Query Match          18.18;   Score 34;   DB 10;   Length 53;
Best Local Similarity  100.0%;   Pred. No. 2.8e-06;
Matches 34;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy     1  GTCGACGGTATCGATAAGCTTGATATCGAATTCG  34
          ||||||||||||||||||||||||||||||||||
Db    43  GTCGACGGTATCGATAAGCTTGATATCGAATTCG  10


RESULT 10
LOCUS       AW172172        54 bp    mRNA    linear     EST 15-NOV-1999
DEFINITION  SWL4CAK07G03SK Brugia malayi L4 cDNA (SAW99MLW-BmL4) Brugia malayi
            cDNA clone SWL4CAK07G03 5', mRNA sequence.
ACCESSION   AW172172
VERSION     AW172172.1  GI:6431968
KEYWORDS    EST.
SOURCE      Brugia malayi
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in fourth stage larvae of Brugia malayi
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Department of Biological Sciences
            Smith College Department of Biological Sciences
            College, Northampton, MA, 01063, USA
            Tel: 413853826
            Fax: 413853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES
    source          Location/Qualifiers
                    1. .54
                    /organism="Brugia malayi"
                    /db_xref="taxon:6279"
                    /clone="SWL4CAK07G03"
                    /clone_lib="Brugia malayi L4 cDNA (SAW99MLW-BmL4)"
                    /dev_stage="larval stage four"
                    /lab_host="XL1-Blue MRF'"
                    /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                    Xho I; Lymphatic filarial nematode parasite of humans.
                    mRNA was prepared from L4s isolated from the peritoneal
                    cavity of jirds and converted to double-stranded cDNA
                    using reverse transcriptase and oligo(dT) followed by
                    RNase H and DNA pol I. The library has 2.7 x 105
```

Column 2

```
                    independent recombinants and the average insert size is
                    approx. 1050bp. The library was constructed by Michelle
                    Lizotte-Wanlewski. The library is available from Dr. S.A.
                    Williams, email:genome@neal.smith.edu."
BASE COUNT         11 a     15 c     11 g     17 t
ORIGIN

Query Match          18.18;   Score 34;   DB 10;   Length 54;
Best Local Similarity  100.0%;   Pred. No. 2.8e-06;
Matches 34;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy     1  GTCGACGGTATCGATAAGCTTGATATCGAATTCG  34
          ||||||||||||||||||||||||||||||||||
Db    43  GTCGACGGTATCGATAAGCTTGATATCGAATTCG  10


RESULT 11
LOCUS       BE758345/c      54 bp    mRNA    linear     EST 16-FEB-2001
DEFINITION  SWYACAL08D01SK Brugia malayi young adult cDNA (SAW99MLW-BmYA)
            Brugia malayi young adult cDNA clone SWYACAL08D01 5', mRNA sequence.
ACCESSION   BE758345
VERSION     BE758345.1  GI:10172337
KEYWORDS    EST.
SOURCE      Brugia malayi
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in young adult of Brugia malayi
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 413853826
            Fax: 413853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES
    source          Location/Qualifiers
                    1. .54
                    /organism="Brugia malayi"
                    /db_xref="taxon:6279"
                    /clone="SWYACAL08D01"
                    /clone_lib="Brugia malayi young adult cDNA (SAW99MLW-BmYA
                    )"
                    /dev_stage="young adult"
                    /lab_host="XL1-Blue MRF'"
                    /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                    Xho I; Lymphatic filarial nematode parasite of humans.
                    mRNA was prepared from young adult worms isolated from the
                    peritoneal cavity of jirds and converted to
                    double-stranded cDNA using reverse transcriptase and
                    oligo(dT) followed by RNase H and DNA pol I. The library
                    has 6.5 x 104 independent recombinants and the average
                    insert size is approx. 800bp. The library was constructed
                    by Michelle Lizotte-Wanlewski. The library is available
                    from Dr. S.A. Williams, email: genome@neal.smith.edu."
BASE COUNT         11 a     15 c     11 g     17 t
ORIGIN

Query Match          18.18;   Score 34;   DB 12;   Length 54;
Best Local Similarity  100.0%;   Pred. No. 2.8e-06;
Matches 34;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy     1  GTCGACGGTATCGATAAGCTTGATATCGAATTCG  34
          ||||||||||||||||||||||||||||||||||
Db    43  GTCGACGGTATCGATAAGCTTGATATCGAATTCG  10


RESULT 12
```

```
LOCUS       AW626545      66 bp    mRNA    linear    EST 30-MAR-2000
DEFINITION  SWOvL3CAN65F04SK Onchocerca volvulus Infective larva cDNA
            (SAW94WL-OVL3) Onchocerca volvulus cDNA clone SWOvL3CAN65F04 5',
            mRNA sequence.
ACCESSION   AW626545
VERSION     AW626545.1  GI:7342410
KEYWORDS    EST.
SOURCE      Onchocerca volvulus
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Williams,S.A., Lu,W., Lizotte-Wanlewski,M. and Laney,S.J.
  TITLE     Genes expressed in infective third stage larvae of Onchocerca
            volvulus
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="Onchocerca volvulus"
                     /strain="Sierra Leone"
                     /db_xref="taxon:6282"
                     /clone="SWOvL3CAN65F04"
                     /clone_lib="Onchocerca volvulus Infective larva cDNA
                     (SAW94WL-OVL3)"
                     /note="Vector: lambda Unizap XR; Site_1: EcoR I; Site_2:
                     Xho I; Cutaneous filarial nematode parasite of humans.
                     mRNA was prepared from third stage infective larvae of
                     Onchocerca volvulus isolated from mosquitoes 10 days after
                     infection and converted to double stranded cDNA using
                     reverse transcriptase and oligo(dT) followed by RNase H
                     and DNApol I. The library had 1.8 x 10E5 independent
                     recombinants and average insert size was 900 base pairs.
                     The library was constructed by Wenhong Lu. The library is
                     available from Dr. S.A. Williams, email genome@smith.edu."
BASE COUNT     11 a     15 c     11 g     29 t
ORIGIN
```

```
RESULT 13
AW159927/c
Query Match            18.1%;  Score 34;  DB 10;  Length 66;
Best Local Similarity  100.0%;  Pred. No. 2.8e-06;
Matches 34;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
         ||||||||||||||||||||||||||||||||||
Db    43 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 10
```

```
LOCUS       AW159927      68 bp    mRNA    linear    EST 05-NOV-1999
DEFINITION  SW14CAK04D11SK Brugia malayi L4 cDNA (SAW99MLW-BmL4) Brugia malayi
            cDNA clone SW14CAK04D11 5', mRNA sequence.
ACCESSION   AW159927
VERSION     AW159927.1  GI:6271956
KEYWORDS    EST.
SOURCE      Brugia malayi
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in fourth stage larvae of Brugia malayi
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1..68
                     /organism="Brugia malayi"
                     /db_xref="taxon:6279"
                     /clone="SW14CAK04D11"
                     /clone_lib="Brugia malayi L4 cDNA (SAW99MLW-BmL4)"
                     /note="Vector: lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Lymphatic filarial nematode parasite of humans.
                     mRNA was prepared from L4s isolated from the peritoneal
                     cavity of jirds and converted to double-stranded cDNA
                     using reverse transcriptase and oligo(dT) followed by
                     RNase H and DNA pol I. The library has 2.7 x 105
                     independent recombinants and the average insert size is
                     approx. 1050bp. The library was constructed by Michelle
                     Lizotte-Waniewski. The library is available from Dr. S.A.
                     Williams, email:genome@neal.smith.edu."
BASE COUNT     14 a     15 c     13 g     26 t
ORIGIN
```

```
RESULT 14
AI066878/c
Query Match            18.1%;  Score 34;  DB 10;  Length 68;
Best Local Similarity  100.0%;  Pred. No. 2.8e-06;
Matches 34;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 34
         ||||||||||||||||||||||||||||||||||
Db    43 GTCGACGGTATCGATAAGCTTGATATCGAATTCG 10
```

```
LOCUS       AI066878      91 bp    mRNA    linear    EST 31-JUL-1998
DEFINITION  SWBmL3SDI03H08SK Brugia malayi L3 subtracted cDNA library
            (SAW97LS-BmL3SD) Brugia malayi cDNA clone SWBmL3SDI03H08 5' or 3',
            mRNA sequence.
ACCESSION   AI066878
VERSION     AI066878.1  GI:3367421
KEYWORDS    EST.
SOURCE      Brugia malayi
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Saunders,L. and Williams,S.A.
  TITLE     Specific and Upregulated Gene Expression in the Third Stage Larvae
            of Brugia malayi
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: SK.
FEATURES             Location/Qualifiers
     source          1..91
                     /organism="Brugia malayi"
                     /db_xref="taxon:6279"
                     /clone="SWBmL3SDI03H08"
                     /clone_lib="Brugia malayi L3 subtracted cDNA library
                     (SAW97LS-BmL3SD)"
```

ORIGIN

Query Match   18.1%;   Score 34;   DB 9;   Length 92;
Best Local Similarity   100.0%;   Pred. No. 2.8e-06;
Matches   34;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG   34
           ||||||||||||||||||||||||||||||||||
Db   41 GTCGACGGTATCGATAAGCTTGATATCGAATTCG   8

Search completed: February 19, 2003, 08:29:42
Job time : 1953 secs

RESULT 15
AA680557/c
LOCUS   AA680557   92 bp   mRNA   linear   EST 04-DEC-1997
DEFINITION SWOvL3CAN03D008SK Onchocerca volvulus Infective larva cDNA
            (SAW94WL-OvL3) Onchocerca volvulus cDNA clone SWOvL3CAN03D008 5',
            mRNA sequence.
ACCESSION   AA680557
VERSION   AA680557.1   GI:2661307
KEYWORDS   EST.
SOURCE   Onchocerca volvulus.
ORGANISM   Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1 (bases 1 to 92)
AUTHORS   Williams,S.A., Lu,W., Lizotte-Wanlewski,M. and Laney,S.J.
TITLE   Genes expressed in infective third stage larvae of Onchocerca
            volvulus
JOURNAL   Unpublished (1995)
COMMENT   Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES   Location/Qualifiers
     source   1..92
                     /organism="Onchocerca volvulus"
                     /strain="Sierra Leone"
                     /db_xref="taxon:6282"
                     /clone="SWOvL3CAN03D008"
                     /clone_lib="Onchocerca volvulus Infective larva cDNA
                     (SAW94WL-OvL3)"
                     /lab_host="Xt1-Blue MRF'"
                     /note="Vector: lambda Unizap XR; Site_1: EcoR I; Site_2:
                     Xho I; Cutaneous filarial nematode parasite of humans.
                     mRNA was prepared from third stage infective larvae of
                     Onchocerca volvulus isolated from mosquitoes 10 days after
                     infection and converted to double stranded cDNA using
                     reverse transcriptase and oligo(dT) followed by RNase H
                     and DNApol I. The library had 1.8 x 10E5 Independent
                     recombinants and average insert size was 900 base pairs.
                     The library was constructed by Wenhong Lu. The library is
                     available from Dr. S.A. Williams, email genome@smith.edu."
BASE COUNT   21 a   26 c   23 g   22 t

/dev_stage="L3 cDNA subtracted with microfilaria, adult
male and adult female cDNA libraries"
/lab_host="Xt1-Blue MRF'"
/note="Vector: Lambda ZAP II; Site_1: Eco RI; Site_2: Eco
RI; Lymphatic filarial nematode parasite of humans. cDNA
from L3, microfilaria, adult male and adult female
libraries was isolated and amplified by PCR using T3 and
T7 primers. Microfilaria, adult male and adult female cDNA
was subtracted from L3 cDNA using the PCR-select cDNA
subtraction kit (Clontech). The library had 4.0 x 105
independent recombinants and the average insert size is
-500bp. The library was constructed by Lori Saunders. The
library is available from Dr. Steven A. Williams, email:
genome@smith.edu."
BASE COUNT   20 a   26 c   23 g   22 t
ORIGIN

Query Match   18.1%;   Score 34;   DB 9;   Length 91;
Best Local Similarity   100.0%;   Pred. No. 2.8e-06;
Matches   34;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 GTCGACGGTATCGATAAGCTTGATATCGAATTCG   34
           ||||||||||||||||||||||||||||||||||
Db   41 GTCGACGGTATCGATAAGCTTGATATCGAATTCG   8